

**UNIVERSIDADE FEDERAL DO RECÔNCAVO DA BAHIA
CENTRO DE CIÊNCIAS AGRÁRIAS, AMBIENTAIS E
BIOLÓGICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS AGRÁRIAS
CURSO DE MESTRADO**

**CARACTERIZAÇÃO E SELEÇÃO DE GENÓTIPOS
BIOFORTIFICADOS DE MANDIOCA E PREDIÇÃO DO
TEOR DE CAROTENOIDES COM BASE EM IMAGENS
DIGITAIS**

RAVENA ROCHA BESSA DE CARVALHO

**CRUZ DAS ALMAS / BAHIA
2021**

CARACTERIZAÇÃO E SELEÇÃO DE GENÓTIPOS BIOFORTIFICADOS DE MANDIOCA E PREDIÇÃO DO TEOR DE CAROTENOIDES COM BASE EM IMAGENS DIGITAIS

Ravena Rocha Bessa de Carvalho
Engenheira Agrônoma
Universidade Federal do Recôncavo da Bahia, 2021

Dissertação apresentada ao Colegiado do Programa de Pós-Graduação em Ciências Agrárias da Universidade Federal do Recôncavo da Bahia, como requisito parcial para obtenção do título de Mestre em Ciências Agrárias (Área de Concentração: Fitotecnia).

Orientador: Prof. Dr. Éder Jorge de Oliveira
Coorientadora: Dra. Luciana Alves de Oliveira
Coorientadora: Dra. Massaine Bandeira e Sousa

**CRUZ DAS ALMAS / BAHIA
2021**

FICHA CATALOGRÁFICA

C331c	<p>Carvalho, Ravena Rocha Bessa de. Caracterização e seleção de genótipos biofortificados de mandioca e predição do teor de carotenoides com base em imagens digitais / Ravena Rocha Bessa de Carvalho._ Cruz das Almas, BA, 2021. 115f.; il.</p> <p>Dissertação (Mestrado) – Universidade Federal do Recôncavo da Bahia, Centro de Ciências Agrárias, Ambientais e Biológicas, Mestrado em Ciências Agrárias. Orientador: Prof. Dr. Éder Jorge de Oliveira. Coorientadora: Prof. Dra. Luciana Alves de Oliveira. Coorientadora: Prof. Dra. Massaine Bandeira e Sousa.</p> <p>1.Mandioca – Variabilidade genética. 2.Mandioca – Melhoramento genético. 3.Germoplasma vegetal – Análise. I.Universidade Federal do Recôncavo da Bahia, Centro de Ciências Agrárias, Ambientais e Biológicas. II.Título.</p> <p>CDD: 633.682</p>
-------	--

Ficha elaborada pela Biblioteca Central de Cruz das Almas -
UFRB.

Responsável pela Elaboração - Antonio Marcos Sarmiento das Chagas (Bibliotecário - CRB5 / 1615). (os dados para catalogação foram enviados pela usuária via formulário eletrônico). ...

**UNIVERSIDADE FEDERAL DO RECÔNCAVO DA BAHIA
CENTRO DE CIÊNCIAS AGRÁRIAS, AMBIENTAIS E
BIOLÓGICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS AGRÁRIAS
CURSO DE MESTRADO**

**CARACTERIZAÇÃO E SELEÇÃO DE GENÓTIPOS
BIOFORTIFICADOS DE MANDIOCA E PREDIÇÃO DO
TEOR DE CAROTENOIDES COM BASE EM IMAGENS
DIGITAIS**

**COMISSÃO EXAMINADORA DA DEFESA DE DISSERTAÇÃO DE
Ravena Rocha Bessa de Carvalho**

Prof. Dr. Éder Jorge de Oliveira
Universidade Federal do Recôncavo da Bahia / UFRB
Examinador Interno (Orientador)

Prof. Dr. Adésio Ferreira
Universidade Federal do Espírito Santo / UFES
Examinador Externo

Prof. Dr. Onildo Nunes de Jesus
Embrapa Mandioca e Fruticultura / CNPMF
Examinador Externo

AGRADECIMENTOS

A Deus por me sustentar até aqui.

A Universidade Federal do Recôncavo da Bahia - UFRB pela oportunidade da realização do curso de Mestrado.

A Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES pela concessão de bolsa de estudo.

A Empresa Brasileira de Pesquisa Agropecuária – Embrapa Mandioca e Fruticultura pela disponibilidade de recursos para o desenvolvimento da minha pesquisa, e a todos os colegas de trabalho.

Ao meu orientador Eder Jorge de Oliveira por todo incentivo e apoio para concretização do curso.

A todos os professores que fazem parte do programa de pós-graduação em Ciências Agrárias, por todos os ensinamentos.

A toda Equipe Mandioca pelo auxílio e parceria no desenvolvimento das atividades. Muito obrigada pela dedicação e amizade de todos.

A minha família por sempre acreditar em mim, em especial meus pais Paulo Sérgio Conceição de Carvalho e Ireulda Rocha Bessa de Carvalho

E a todos que torcem pelas minhas conquistas e me auxiliaram nesse processo, mesmo de longe.

Entrega o teu caminho ao Senhor, confia nele, e o mais ele fará.
Salmos 37:5

SUMÁRIO

	Página
Revisão de literatura.....	1
Referências Bibliográficas	19

Artigo I

DIVERSIDADE FENOTÍPICA E SELEÇÃO NO GERMOPLASMA BIOFORTIFICADO DE MANDIOCA PARA CARACTERÍSTICAS PRODUTIVAS E DE QUALIDADE DA RAIZ.....	30
--	----

Artigo II

FENOTIPAGEM BASEADA EM IMAGENS DIGITAIS DE RAÍZES DE MANDIOCA PARA ESTUDOS DE DIVERSIDADE E PREDIÇÃO DO TEOR DE CAROTENOIDES.....	80
---	----

CARACTERIZAÇÃO E SELEÇÃO DE GENÓTIPOS BIOFORTIFICADOS DE MANDIOCA E PREDIÇÃO DO TEOR DE CAROTENOIDES COM BASE EM IMAGENS DIGITAIS

Autora: Ravena Rocha Bessa de Carvalho

Orientador: Dr. Eder Jorge de Oliveira

RESUMO: O desenvolvimento de variedades de mandioca (*Manihot esculenta* Crantz) biofortificadas (alto teor de carotenoides) depende da existência de variabilidade genética e de métodos precisos e otimizados de seleção. Este trabalho avaliou a variabilidade genética para teor de carotenoides totais (TCC) e atributos agronômicos de produtividade e qualidade de raízes, bem como o uso de imagens digitais para a predição do teor de carotenoides totais em genótipos de mandioca. A variabilidade genética para TCC foi bastante elevada ($0,075$ a $13,08 \mu\text{g}\cdot\text{g}^{-1}$), assim como para as demais características avaliadas, como consequência, a análise deviance indicou significância do efeito de genótipos para todas as características. A herdabilidade no sentido amplo para TCC foi de alta magnitude ($h^2=0,72$) e de baixa a média magnitude (h^2 entre $0,21$ e $0,60$) para as demais características. O TCC possui forte correlação com a cor de polpa ($r=0,70$), mas praticamente nula e sem significância para teor de matéria seca nas raízes. A seleção dos 30 genótipos para recombinação no programa de melhoramento de mandioca tem potencial para elevar o TCC em 37% e reduzir o teor de compostos cianogênicos em -19,4%. Além disso, foi analisado o potencial de uso de imagens digitais para predição do TCC em raízes de mandioca, com base nos índices L^* , a^* e b^* , Hue e Croma do sistema de cores CIE (*International Commission on Illumination*) e luminosidade. Alta correlação positiva entre TCC e os índices colorimétricos b^* ($r=0,90$) e croma ($r=0,89$) foi observada. Dentre os diferentes modelos de predição analisados, o *artificial neural network* apresentou a maior capacidade preditiva ($R^2=0,94$), associada ao menor erro nas estimativas do TCC (RMSE= 0,24). A estrutura da população estudada revelou cinco grupos com base na similaridade quanto aos índices colorimétricos e ao teor de carotenoides totais, sendo possível a formação de grupos bastante contrastantes para estas características.

Palavras-chave: *Manihot esculenta* Crantz, germoplasma, predição, qualidade de raízes.

CHARACTERIZATION AND SELECTION OF BIOFORTIFIED CASSAVA GENOTYPES AND PREDICTION OF CAROTENOID CONTENT BASED ON DIGITAL IMAGES

Author: Ravena Rocha Bessa de Carvalho
Adivisor: Dr. Eder Jorge de Oliveira

ABSTRACT: The development of biofortified cassava (*Manihot esculenta* Crantz) varieties (high carotenoids content) depends on the genetic variability, as well as of precise and optimized methods for screening a large number of genotypes. The objectives of this work were to study the genetic variability for total carotenoid content (TCC), root yield and root quality attributes, as well as the use of digital images to predict TCC in cassava genotypes. The genetic variability for TCC (range of 0.075 to 13.08 $\mu\text{g.g}^{-1}$) as well as other traits evaluated was quite high; as a consequence the deviance analysis indicated the significance effect of genotypes for all traits. The broad-sense heritability for TCC had high magnitude ($h^2=0.72$), and low to medium magnitude (h^2 between 0.21 and 0.60) for the other agronomic and root quality traits. TCC was strongly correlated with pulp color ($r=0.70$), but practically null with no significance for dry matter content in the roots. The selection of the 30 genotypes for recombination in the cassava breeding program has the potential to increase TCC by 37% and reduce the content of cyanogenic compounds by -19.4%. In addition, the potential for using digital images to predict TCC in cassava roots was analyzed, based on the L^* , a^* e b^* , Hue and Chroma indexes of the CIE system (International Commission on Illumination) and luminosity. A high positive correlation between TCC and the colorimetric indices b^* ($r=0.90$) and chroma ($r=0.89$) was observed. Among the different prediction models analyzed, the artificial neural network showed the highest predictive ability ($R^2=0.94$), associated with the smallest error of the TCC estimates (RMSE=0.24). Regarding the population structure we identified five groups based on similarity in colorimetric indices and total carotenoid content, making it possible the identification of very contrasting groups for these characteristics.

Keywords: *Manihot esculenta* Crantz, germplasm, prediction, root quality.

REFERENCIAL TEÓRICO

1. Aspectos gerais da cultura da mandioca

A mandioca (*Manihot esculenta* Crantz) é uma cultura perene, pertencente à família Euphorbiaceae. A mandioca é alógama (polinização cruzada) com flores masculinas e femininas presentes na mesma planta (monóica). O fruto apresenta formato de uma cápsula e as sementes são elipsoidais (IITA, 1990; OSPINA; CEBALLOS, 2012). Ela possui $2n = 36$ cromossomos, e é considerado um anfidiplóide ou alopoliplóide sequencial (EL-SHARKAWY, 2003).

A propagação da mandioca é feita predominantemente de forma assexuada, mas a forma sexuada também pode ocorrer, sobretudo quando induzida para fins de melhoramento genético (OSPINA; CEBALLOS, 2012). Comercialmente, a propagação utilizada é assexuada, cujas estacas do caule são conhecidas popularmente por manivas e estas contribuem para o sucesso da propagação e estabelecimento da cultura em campo, uma vez que os aspectos fitossanitários sejam observados, a fim de evitar disseminação de pragas e doenças (ESCOBAR, et al., 2006). As principais vantagens da propagação assexuada é a capacidade de assegurar que a planta selecionada repasse integralmente suas características genéticas para seus clones, garantindo que a nova população seja agronomicamente equivalente à planta mãe. Por outro lado, a desvantagem da propagação assexuada se deve ao maior cuidado necessário para manutenção da qualidade genética e fitossanitária ao longo dos diferentes anos de cultivo.

A propagação sexuada por meio de sementes botânicas não é viável comercialmente, pela geração de variabilidade genética e consequente heterogeneidade das plantas no campo, considerando que a mandioca é uma planta alógama e altamente heterozigótica. Entretanto, possibilita a geração de diversidade genética para o desenvolvimento de variedades melhoradas em função da recombinação alélica entre parentais contrastantes (RAMALHO; SANTOS; PINTO, 2008). Algumas características importantes para o desenvolvimento de novas variedades melhoradas geneticamente são a alta produtividade de raízes, resistência a pragas e doenças (OLIVEIRA; OLIVEIRA

FILHO; SANTOS, 2015), menor toxicidade (PARMAR; STURM; HENSEL, 2017), tolerância ao déficit hídrico (KOUNDINYA et al., 2018) e melhor conteúdo nutricional (ESUMA et al., 2012; CEBALLOS et al., 2013).

A raiz da mandioca é o órgão de maior interesse, pois, devido sua constituição ser basicamente carboidratos, mais precisamente o amido, torna-se uma fonte energética de alto valor para diversos fins. O maior consumo das raízes de mandioca ocorre em regiões tropicais, tanto em preparos simples como o cozimento, quanto em produtos feitos a partir do seu processamento (FAO, 2013; FAO, 2018). Apesar disso, a mandioca é uma planta em que se aproveita todas as suas partes (folhagem, caules e raízes) (CEBALLOS et al., 2017). De modo geral, a cultura da mandioca é de fácil cultivo, têm bons rendimentos em solos pouco férteis mesmo na ausência de fertilizantes, e ainda apresenta certa tolerância ao déficit hídrico (JARVIS et al., 2012). As raízes são perecíveis em razão do alto conteúdo de umidade, e o início da deterioração pós-colheita acontece em um tempo relativamente curto, a qual se dá por processos enzimáticos ou microbiológicos (RAHMAWATI et al., 2021).

2. Importância socioeconômica da cultura

A mandioca é a principal cultura alimentar da África e por esse motivo o continente africano se caracteriza como principal produtor e consumidor. Nesse continente, a mandioca é uma escolha estratégica no contexto de adaptações às mudanças climáticas. Na Ásia (segunda maior região produtora mundial), a demanda industrial da mandioca para utilização do amido, ração animal e etanol, bem como as suas lucrativas exportações, especialmente para a China, sustentaram a forte expansão da safra nos últimos anos. Por outro lado, na América Latina, existe grande expectativa de aumento de produção, especialmente no Brasil, onde tem havido importantes movimentos para incentivar a combinação amido de trigo x mandioca na produção de massas (FAO, 2018).

Em nível mundial, em 2019 a Nigéria assumiu a liderança em produção de raízes de mandioca, seguida pela República Democrática do Congo, Tailândia e Gana, com uma produção de 59,19; 40,05; 31,07; e 22,44 milhões de toneladas de raízes, respectivamente (FAO, 2019). O Brasil se destaca

como o 5º maior produtor mundial, com produção de 17,49 milhões de toneladas, apresentando área colhida de 1,19 milhões de hectares, e um rendimento médio de 14,70 t.ha⁻¹ (FAO, 2019). Dentre os estados brasileiros, o Pará (com 3,7 milhões de toneladas) se destaca com maior produção, seguido de Paraná (com 3,2 milhões de toneladas) e São Paulo (com 1,4 milhões de toneladas). A Bahia está em sétimo lugar no ranking de maior produtora (com 0,65 milhão de toneladas), sendo o estado com a maior produção na região Nordeste, e terceiro lugar no ranking de área colhida (com 92.788 hectares) (IBGE, 2019).

Na região Norte do Brasil, existe o consumo tradicional de diversos produtos derivados da raiz e da folha da mandioca. Os principais produtos obtidos das raízes da mandioca são: farinhas d'água, seca e mista (mistura da massa ralada e fermentada); goma, subproduto da farinha, que é utilizada para consumo familiar e comercialização para fabricação da farinha de tapioca, biscoitos e o tucupi (molho feito com manipueira e pimenta) que é extraído de raízes amarelas e participa da confecção de pratos típicos da região, a exemplo do Tacacá (caldo do tucupi com folhas de jambú) (CARDOSO et al., 2001; SONATI; VILARTA; SILVA, 2009).

A forma mais comum de consumir as folhas de mandioca é no prato típico conhecido como maniçoba, que leva uma mistura de carnes e folhas de mandioca cozidas. Na região Nordeste é a farinha de mandioca que se destaca entre os produtos derivados dessa raiz (SONATI; VILARTA; SILVA, 2009). Quanto à questão nutricional das folhas, na matéria seca é possível que se obtenha cerca de 30% de proteína bruta e grandes teores de minerais, tais como, zinco, ferro, cálcio, fósforo, potássio e magnésio e vitaminas (B1, B2, C e carotenoides) (LATIF; MÜLLER, 2015).

As folhas são comumente consumidas, além do Brasil, em outros países da América do Sul e em vários países da África e Ásia (PARMAR; STURM; HENSEL, 2017). O Congo foi listado como maior consumidor de folhas de mandioca, constituindo cerca de 60% de todo o consumo de vegetais (LATIF; MÜLLER, 2015). Na Serra Leoa e na Libéria também são consumidas pelas mulheres grávidas para aumentar a produção de leite materno (AREGHEORE, 2012).

A mandioca é uma cultura com importante papel na segurança alimentar de populações menos favorecidas economicamente, além de ser matéria prima em uma vasta gama de produtos processados importantes para o crescimento dos mercados nacionais e para as exportações, dentre eles estão: tapioca, farinhas, amidos modificados, *chips*, bebidas, rações para alimentação animal, ácido láctico, glicose líquida, frutose, maltose, plásticos biodegradáveis e produção de bioetanol (BALAGOPALAN, 2002; LU; DING; WU, 2011; JIMÉNEZ et al., 2012; HEUZÉ et. al., 2016; FAO, 2018).

Toda a planta pode ser utilizada para alimentação animal, por meio de diferentes processamentos, a exemplo da trituração, secagem e ensilagem, os quais promovem a redução do teor de compostos cianogênicos e aumentam a qualidade nutricional (ORORORO et al., 2014; LI et al., 2019). As folhas são uma ótima fonte de proteínas para os animais, com níveis variando de 16,6% a 39,9% (KHIEU et al., 2005) e no caso das vacas leiteiras, a suplementação com 1 a 2 kg de feno de mandioca por animal por dia, garante a melhoria do desempenho (FAO, 2013).

3. Variabilidade genética para características relacionadas à qualidade de raiz e componentes de produtividade em mandioca

A mandioca tem sua origem na região Sul da Amazônia (OLSEN; SCHAAL, 1999) e segundo estudo de Léotard et al. (2009), sua domesticação ocorreu apenas uma vez, mais precisamente no sudoeste da Amazônia, que atualmente corresponde aos estados de Rondônia e Mato Grosso. São nesses dois estados, juntamente com Tocantins, Goiás e Acre, que ocorrem as populações silvestres dessa cultura no Brasil (OLSEN; SCHAAL, 1999).

Para assegurar a manutenção da variabilidade genética das plantas foram criados os bancos de germoplasma, em condições seminaturais, o que garante a disponibilidade de genes para o melhoramento genético (RABBI et al., 2017). Algumas organizações internacionais e nacionais têm atuado na conservação de germoplasma e no melhoramento de cultivares de mandioca a exemplo do Instituto Internacional de Agricultura Tropical (IITA), Ibadan, Nigéria; Centro Internacional de Agricultura Tropical (CIAT), Cali, Colômbia; Embrapa Mandioca e Fruticultura (CNPMPF), Bahia, Brasil; Instituto Central de

Pesquisa de Tubérculos (CTCRI), Índia; Centro de Pesquisa de Culturas de Campo Rayong, Tailândia e institutos similares na Indonésia, Filipinas e Vietnã (PARMAR; STURM; HENSEL, 2017).

A conservação da variabilidade genética do Brasil tem sido feita de forma *in situ* e *ex situ*, realizada pela EMBRAPA Recursos genéticos e Biotecnologia (CENARGEN), que iniciou as coletas em 1975 e em 1993 já contava com mais de 50.000 acessos de diferentes culturas, sendo que algumas espécies de propagação vegetativa, como a mandioca, foram conservadas *in vitro* (RAMALHO; SANTOS; PINTO, 2008). A Embrapa Mandioca e Fruticultura (Cruz das Almas, Bahia, Brasil), mantém mais de 1.600 acessos *in vivo*. Essa variabilidade é representada principalmente por variedades selecionadas naturalmente pelos produtores e por variedades melhoradas (OLIVEIRA et al., 2014).

A avaliação da diversidade genética e caracterização do germoplasma de mandioca é uma importante estratégia, por contribuir com informações relacionadas à variabilidade genética das cultivares e também identificar combinações parentais que permitam o desenvolvimento de progênieis segregantes com máxima variabilidade genética e introgressão de genes desejáveis presentes no germoplasma disponível (OLIVEIRA et al., 2014). Para isto, a análise simultânea de diferentes variáveis pode proporcionar uma melhor representação da real variabilidade genética (OLIVEIRA; OLIVEIRA FILHO; SANTOS, 2015).

A produtividade de raízes de mandioca depende de fatores externos relacionados ao ambiente e ao manejo cultural, embora o componente genético seja o mais determinante na obtenção de altas produtividades. A produtividade de raízes é uma das principais características a serem buscadas na seleção de genótipos de mandioca, pois está intimamente relacionada à rentabilidade da cultura (FUHRMAN et al., 2019). Entretanto, diversas outras características de desempenho agrônomo devem ser levadas em consideração para o sucesso de uma variedade (OLIVEIRA; OLIVEIRA FILHO; SANTOS, 2015; FUHRMANN et al., 2019).

A fim de suprir uma dieta deficiente em nutrientes, estratégias como a biofortificação, buscam a obtenção de variedades com alto conteúdo de

carotenoides (SAYRE et al., 2011). Por muito tempo, o projeto global HarvestPlus, objetivou promover o desenvolvimento de variedades com maiores teores de vitaminas e minerais por meio de métodos convencionais de melhoramento. A principal estratégia envolveu a identificação de variedades com altos teores de micronutrientes e posterior cruzamento com variedades com características agrônômicas desejáveis (SALTZMAN et al., 2017), e desde então o aumento do teor de carotenoides em raízes de mandioca tem sido foco constante em programas de melhoramento da espécie (ESUMA et al., 2012; CEBALLOS et al., 2013; RABBI et al., 2017; CEBALLOS et al., 2017; BEYENE et al., 2017). Neste caso, o foco é o desenvolvimento de variedades de mesa para consumo cozida ou minimamente processada, que deve possuir baixo conteúdo de compostos cianogênicos.

Toda a planta da mandioca apresenta toxicidade, devido à presença de glicosídeos cianogênicos, sobretudo a linamarina (95%) e lotaustralina (5%), que são sintetizados a partir dos aminoácidos valina e isoleucina, respectivamente (RIJSSEN; MORRIS; ELOFF, 2013). Estes compostos são defesas naturais da planta contra insetos e animais e a concentração desses compostos podem variar entre diferentes cultivares, a depender da disponibilidade de nutrientes do solo e das condições ecológicas em que a cultura está se desenvolvendo (PARMAR; STURM; HENSEL, 2017). A enzima linamarase está presente na parede celular e os compostos cianogênicos estão compartimentalizados dentro da célula. Uma vez ocorrendo a ruptura nas células da planta por meio de danos mecânicos, a linamarina entra em contato com a linamarase, provocando a liberação do ácido cianídrico (HCN) (PARMAR; STURM; HENSEL, 2017). No Brasil, as variedades de mandioca são denominadas “mansas” ou “bravas”. As bravas contêm teores a partir de 100 mg de ácido cianídrico/kg de polpa de acordo com o conteúdo de glicosídeos cianogênicos (FRANCK et al., 2011).

O teor de glicosídeos cianogênicos das folhas é superior em relação às raízes; entretanto são degradados durante o cozimento ou processamento de forma semelhante ao que ocorre nas raízes (LATIF; MÜLLER, 2015; PARMAR; STURM; HENSEL, 2017). A ingestão de grandes quantidades de cianeto por humanos pode levar a intoxicação causando diversos sintomas, como dores de

cabeça, náuseas, tonturas, diarreia, vômitos e em casos extremos, a morte (NHASSICO et al., 2008). Por isso, é importante que o melhoramento genético desenvolva variedades com teores de compostos cianogênicos em níveis não prejudiciais à saúde humana (PARMAR; STURM; HENSEL, 2017).

O amido é o principal componente presente nas raízes e ainda o principal produto explorado da mandioca, chegando a cerca de 80% do total da matéria seca (SÁNCHEZ et al., 2009). De modo geral, o acúmulo máximo da matéria seca acontece no final do período seco, pouco antes da chegada das chuvas (CEBALLOS; HERSHEY; BECERRA-LÓPEZ-LAVALLE, 2012). O teor de matéria seca nessas raízes varia em torno de 30% (em matéria fresca), podendo alcançar maiores amplitudes em função dos ganhos genéticos em programas de melhoramento (SÁNCHEZ et al., 2009; SÁNCHEZ et al., 2014; RABBI et al., 2017).

Devido à alta correlação entre o conteúdo de matéria seca e o amido, o melhoramento genético tem buscado aproveitar esta correlação positiva para seu aumento simultâneo. Por esse motivo, estudos recentes tem como foco a determinação do conteúdo de matéria seca de forma mais eficiente e acurada, além da compreensão das causas genéticas que interferem no seu acúmulo ao longo do ciclo de desenvolvimento da planta (SÁNCHEZ et al., 2014; RABBI et al., 2017; IKEOGU et al., 2017).

O amido é constituído por dois tipos de polímeros de glicose: a amilose e a amilopectina (LEMOS et al., 2019). Baseado em uma grande quantidade de genótipos, uma estimativa bastante robusta do teor médio de amilose é de 20,7% (SÁNCHEZ et al., 2009). O conteúdo desse polímero é um parâmetro de qualidade importante para produtos à base de amido, pois afeta propriedades como gelatinização, solubilidade, característica de pasta e textura (ZHU et al., 2008).

4. Carotenoides

Os carotenoides são pigmentos responsáveis pelas cores que variam do amarelo ao vermelho em frutos, vegetais, fungos e flores (MALDONADO-ROBLEDO et al., 2003) e desempenham papel fundamental na saúde humana, a exemplo do β -caroteno, que dependendo de sua metabolização atua como

antioxidante, extinguindo o oxigênio singlete e eliminando radicais peroxil (STROBEL; TINZ; BIESALSKI, 2007).

O β -caroteno é o carotenoide encontrado em maior quantidade na mandioca e atua como precursor da vitamina A. Ele é absorvível, biodisponível e convertido em retinol no corpo humano (CEBALLOS et al., 2017). Na fotossíntese, os carotenoides têm função de pigmento para absorção de luz e fotoprotetores contra danos oxidativos (DELGADO-VARGAS; JIMÉNEZ; PAREDES-LÓPES, 2000).

O β -caroteno e o licopeno são exemplos de carotenos que são carotenoides hidrocarbonetos, já os carotenoides com funções químicas oxigenadas são chamados de xantofilas, como a β -criptoxantina, cantaxantina. Além disso, eles podem ser acíclicos (licopeno), monocíclicos (γ -caroteno) ou bicíclicos (α - e β -caroteno). Se apresentam geralmente na forma *trans*, embora haja a forma *cis* (RODRIGUEZ-AMAYA; KIMURA; AMAYA-FARFAN, 2008).

A biofortificação da mandioca é um processo que busca o aumento do conteúdo de β -caroteno por meio de técnicas convencionais de melhoramento e pela transformação genética, a fim de obter variedades enriquecidas nutricionalmente e introduzi-las em áreas onde há o consumo mais intenso da cultura, visando à redução da deficiência de vitamina A (OLURANTI; BADEJO; FAGBEMI, 2016; SALTZMAN et al., 2017; CEBALLOS et al., 2017).

A existência de variabilidade genética para teor de carotenoides tem atraído a atenção dos melhoristas de mandioca na tentativa de desenvolver variedades com características agronômicas adequadas ao cultivo, e que ao mesmo tempo sejam biofortificadas e com isso possam agregar valor nutricional à dieta (SAYRE et al., 2011; CEBALLOS et al., 2017; FUHRMANN et al., 2019). Nos últimos anos, o programa de melhoramento do Centro Internacional de Agricultura Tropical (CIAT) praticamente triplicou a concentração original de carotenoides nas raízes de mandioca por meio do melhoramento convencional (CEBALLOS; HERSHEY; BECERRA-LÓPEZ-LAVALLE, 2012; CEBALLOS et al., 2013; SÁNCHEZ et al., 2014). Este progresso foi observado ao longo dos ciclos de seleção, apresentando valor máximo de $10,3 \mu\text{g.g}^{-1}$ no ano de 2004 e $24,3 \mu\text{g.g}^{-1}$ no ano de 2012 (CEBALLOS et al., 2013).

Os pigmentos carotenoides são degradados pelo calor, luz e oxigênio (RODRIGUEZ-AMAYA; KIMURA; AMAYA-FARFAN, 2008). Maziya-Dixon; Dixon; Ssemakula (2009) constataram que o conteúdo de carotenoides total foi reduzido quando as raízes cruas foram descascadas e raladas em todas as variedades incluídas no estudo, e essa redução possivelmente foi resultado da oxidação devido a exposição a luz e ao ar. Além disso, o processamento quebra a matriz da planta, incluindo os compartimentos celulares e a ligação de proteínas que servem para proteger e estabilizar os carotenoides (OLURANTI; BADEJO; FAGBEMI, 2016).

5. Fenotipagem de alto rendimento por meio de imagens

A fenotipagem é descrita como um conjunto de metodologias e protocolos usados para medir o crescimento, a arquitetura e a composição das plantas com exatidão e precisão em diferentes escalas de organização, de órgãos a dosséis. Seu objetivo é fornecer dados quantitativos sobre as respostas dinâmicas entre as plantas e o meio ambiente (FIORANI; SCHURR, 2013).

A imagem digital surgiu como um marco para a captura de informações fenotípicas quantitativas na maioria das abordagens automatizadas ou semiautomáticas. A geração de imagens permitiu que muitos aspectos do desenvolvimento, função e saúde da planta fossem monitorados, medidos e rastreados de maneiras anteriormente inatingíveis usando métricas convencionais (COBB et al., 2013). Essa abordagem como alternativa de fenotipagem de alto rendimento tem se tornando cada vez mais comum em diversas linhas de pesquisa (FIORANI; SCHURR, 2013).

A utilização de imagens para avaliar características baseadas em cores apresenta fortes correlações entre dados digitais e visuais. Esse método de fenotipagem resulta em economia de tempo de análise, permite utilizar câmeras digitais disponíveis comercialmente e softwares de código aberto, demonstrando alto potencial para uso em programas de melhoramento de plantas (WALTER et al., 2019). Esta ferramenta permite extrair informações de cor das imagens, correlacionar com a característica associada e a partir daí

desenvolver modelos para quantificar indiretamente a característica (BAI et al., 2018).

Entretanto, a medição de cores de imagens digitais requer padronização e interpretação porque os dispositivos digitais usam o espaço de cores Red Green Blue (RGB) que não é padronizado, não é linear e pode variar entre dispositivos de hardware e aplicativos de software (DARRIGUES et al., 2008). O software *Tomato Analyzer* é gratuito e dispõe da ferramenta *Color Test* para coletar medições objetivas de cores de imagens JPEG, de forma confiável, precisa e acessível. Essa ferramenta registra os valores RGB de cada pixel do objeto selecionado e os traduz em valores médios L^* , a^* e b^* do espaço de cores CIELAB (CIE, 1978), que é um padrão de referência comumente utilizado para medir a cor do objeto (DARRIGUES et al., 2008). O valor de L^* define a luminosidade ($L^* = 0$ escuridão e $L^* = 100$ claridade) e a^* e b^* são responsáveis pela cromaticidade ($+a^*$ vermelho e $-a^*$ verde), ($+b^*$ amarelo e $-b^*$ azul). Por meio do módulo L^* , a^* e b^* , é possível calcular o chroma (chroma ou intensidade de cor; 0 - cor impura e 60 - cor pura) por meio da fórmula $[(a^{*2} + b^{*2})^{1/2}]$, e o hue (tonalidade ou ângulo de cor; 0° vermelho; 90° amarelo; 180° verde; 270° azul e 360° negro) por meio das fórmulas $H^\circ = \tan^{-1} \left(\frac{b^*}{a^*} \right)$, quando $a^* > 0$ e $b^* \geq 0$, e $H^\circ = 180 + \tan^{-1} \left(\frac{b^*}{a^*} \right)$, quando $a^* < 0$.

Diferentes estudos utilizaram imagens para avaliar características baseadas em cor, utilizando diferentes dispositivos de captura, em diferentes culturas, como detecção de tomates maduras (cor vermelha) (KHOSHROO; AREFI; KHODAEI, 2014), detecção de frutos imaturos de pêssego (KURTULMUS; LEE; VARDAR, 2014), desenvolvimento de pontuação automatizada de clorose por deficiência de ferro em soja (BAI et al., 2018), para investigar as respostas morfofisiológicas do milho ao suprimento de nitrogênio e água (DODIG et al., 2019), identificação do amarelecimento fisiológico do trigo (WALTER et al., 2019) e para caracterização e compreensão da diversidade de plantas (LI; FRANK; MIGICOVSKY, 2019).

Na cultura da mandioca, essa ferramenta também tem demonstrado excelentes resultados, em estudo de associação de características de arquitetura de raiz da mandioca com concentração de micronutrientes

(BUSENER et al., 2020), estimativa do crescimento e do desempenho nutricional da mandioca sob irrigação deficitária e fertirrigação com potássio (WASONGA et al., 2021) e avaliação de características relacionadas à estrutura do dossel e rendimento da raiz (SELVARAJ et al., 2020).

Em relação à estimativa indireta do teor de carotenoides por meio da fenotipagem baseada em imagens, resultados promissores foram descritos por Ye; Izawa; Zhang (2018), que desenvolveram um aplicativo para determinação rápida do teor de licopeno nos frutos de tomate e classificação com base nas cores, por González et al. (2020), que apresentaram uma metodologia para predição do teor de carotenoides individuais a partir de parâmetros de análise de imagens digitais de amostras de pólen de espécies das famílias Brassicaceae, Myrtaceae, Fabaceae, e por Aditya; Saputro; Handayani (2018), que implementaram um algoritmo genético no sistema de predição de imagens Visible-NIR para predição do conteúdo de carotenoides em *Amaranthus* sp, O estudo da relação entre índices de cor extraídos das imagens e teor de carotenoides foram demonstrados na cultura da cenoura (*Daucus carota*) (PACE et al., 2013), tomate (*Solanum lycopersicum* L.) (STINCO et al., 2013), azeite de oliva (*Olea europaea*) (SANAEIFAR; JAFARI; GOLMAKANI, 2018) e tangerina (*Citrus reticulata* Blanco) (HADIMAN; MITTAL, 2019).

REFERÊNCIAS BIBLIOGRÁFICAS

ADITYA, K.; SAPUTRO, A. H; HANDAYANI, W. Enhancement of visible-NIR imaging prediction system using genetic algorithm: prediction of carotenoid content in *Amaranthus* sp. Leaf. In: **2018 International Conference on Electrical Engineering and Informatics (ICELTICs)**. IEEE, 2018. p. 106-110.

AREGHEORE, E. M. Nutritive value and inherent anti-nutritive factors in four indigenous edible leafy vegetables in human nutrition in Nigeria: a review. **Journal of Food Resource Science**, v. 1, p. 1-14, 2012.

BAI, G.; JENKINS, S.; YUAN, W.; GRAEF, G. L.; GE, Y. Field-based scoring of soybean iron deficiency chlorosis using RGB imaging and statistical learning. **Frontiers in Plant Science**, v. 9, p. 1002, 2018.

BALAGOPALAN, C. Cassava utilization in food, feed and industry. **Cassava: Biology, production and utilization**, 2002, 332p.

BEYENE, G.; SOLOMON, F. R.; CHAUHAN, R. D.; GAITÁN-SOLIS, E.; NARAYANAN, N.; GEHAN, J.; SIRITUNGA, D.; STEVENS, R. L.; JIFON, J.; ECK, J. V.; LINSLER, E.; GEHAN, M.; ILYAS, M.; FREGENE, M.; SAYRE, R. T.; ANDERSON, P.; TAYLOR, N. J.; CAHOON, E. B. Provitamin A biofortification of cassava enhances shelf life but reduces dry matter content of storage roots due to altered carbon partitioning into starch. **Plant biotechnology journal**, v. 16, n. 6, p. 1186-1200, 2017.

BUSENER, N.; KENGGANNA, J.; SAENGWILAI, P. J.; BUCKSCH, A. Image-based root phenotyping links root architecture to micronutrient concentration in cassava. **Plants People Planet**, v. 2, n. 6, p. 678-687, 2020.

CARDOSO, E. M. R.; MÜLLER, A. A.; SANTOS, A. I. M. DOS; HOMMA, A. K. O.; ALVES, R. N. B. Processamento e comercialização de produtos derivados da mandioca no nordeste paraense. **Embrapa Amazônia Oriental-Documentos (INFOTECA-E)**, 2001.

CEBALLOS, H.; DAVRIEUX, F.; TALSMA, E. F.; BELALCAZAR, J.; CHAVARRIAGA, P.; ANDERSSON, M. S. Carotenoids in cassava roots. In: **Carotenoids**. InTech, p. 189-221, 2017.

CEBALLOS, H.; HERSHEY, C.; BECERRA-LÓPEZ-LAVALLE, L. A. New approaches to cassava breeding. **Plant Breeding Reviews**, v. 36, p. 427–504, 2012.

CEBALLOS, H.; MORANTE, N.; SÁNCHEZ, T.; ORTIZ, D.; ARAGÓN, I.; CHÁVEZ, A. L.; PIZARRO, M.; CALLE, F.; DUFOUR, D. Rapid cycling recurrent selection for increased carotenoids content in cassava roots. **Crop Science**, v. 53, n. 6, p. 2342-2351, 2013.

COBB, J. N.; DECLERCK, G.; GREENBERG, A.; CLARK, R.; MCCOUCH, S. Next-generation phenotyping: requirements and strategies for enhancing our understanding of genotype–phenotype relationships and its relevance to crop improvement. **Theoretical and Applied Genetics**, v. 126, n. 4, p. 867-887, 2013.

Commission Internationale de l'Éclairage (CIE). Recommendations on uniform color spaces, color-difference equations, psychometric color terms. **Paris: CIE**, 1978.

DARRIGUES, A.; HALL, J.; VAN DER KNAAP, E.; FRANCIS, D. M.; DUJMOVIC, N.; GRAY, S. Tomato analyzer-color test: a new tool for efficient digital phenotyping. **Journal of the American Society for Horticultural Science**, v. 133, n. 4, p. 579-586, 2008.

DELGADO-VARGAS, F.; JIMÉNEZ, A. R.; PAREDES-LÓPES, O. Natural pigments: carotenoids, anthocyanins, and betalains—characteristics, biosynthesis, processing, and stability. **Critical reviews in food science and nutrition**, v. 40, n. 3, p. 173-289, 2000.

DODIG, D.; BOŽINOVIC, S.; NIKOLIC, A.; ZORIC, M.; VANCETOVIC, J.; IGNJATOVIC-MICIC, D.; DELIC, N.; WEIGELT-FISCHER, K.; JUNKER, A.; ALTMANN, T. Image-derived traits related to mid-season growth performance of maize under nitrogen and water stress. **Frontiers in plant science**, v. 10, p. 814, 2019.

EL-SHARKAWY, M. A. Cassava biology and physiology. **Plant molecular biology**, v. 53, n. 5, p. 621-641, 2003.

ESCOBAR, R. H.; R. H., ANDREZ, C. H.; LARRAHONDO, N.; OSPINA, G.; RESTREPO, J.; NOZ, L. M; ROCA, W. M. Tissue culture for farmers:

Participatory adaptation of low-input cassava propagation in Colombia. **Experimental agriculture**, v. 42, n. 1, p. 103-120, 2006.

ESUMA, W.; RUBAIHAYO, P.; PARIYO, A.; KAWUKI, R.; WANJALA, B.; NZUKI, I.; HARVEY, J. J. W.; BAGUMA, Y. Genetic diversity of provitamin A cassava in Uganda. **Journal of Plant Studies**, v. 1, n. 1, p. 60, 2012.

FIORANI, F.; SCHURR, U. Future Scenarios for Plant Phenotyping. **Annual Review of Plant Biology**, v. 64, n. 1, p. 267-291, 2013.

FOUNDATION AGRICULTURAL ORGANIZATION (FAO). Food Outlook: **BIANNUAL REPORT ON GLOBAL FOOD MARKETS**, 2018. Disponível em: <<http://www.fao.org/3/CA2320EN/ca2320en.pdf>>. Acesso em: 10 de Março de 2021.

FOUNDATION AGRICULTURAL ORGANIZATION (FAO). **Participação dos continentes na produção de mandioca em 2019**. FAOSTAT Database Gateway – FAO. Disponível em: <<http://www.fao.org/faostat/en/#data/QC>>. Acesso em: 18 de Maio de 2021.

FOUNDATION AGRICULTURAL ORGANIZATION (FAO). **Produzir mais com menos: Mandioca**. Um guia para a intensificação sustentável da produção. Informe de Política. 2013. Acesso em: 20 novembro de 2018.

FRANCK, H.; MESTRES, C.; NOËL, A.; BRIGITTE, P.; JOSEPH, H.D.; CORNET, D.; MATHURIN, N.C. Effects of cultivar and harvesting conditions (age, season) on the texture and taste of boiled cassava root. **Food Chemistry**, v.126, p.127-133, 2011.

FUHRMANN, E.; VIEIRA, E. A.; FIALHO, J. DE F.; FALEIRO, F. G.; CARVALHO, L. J. C. B. Agronomic performance and biochemical attributes of yellow-pulped elite sweet cassava clones. **Científica**, v. 47, n. 1, p. 77-82, 2019.

GONZÁLEZ, C. S.; RODRÍGUEZ-PULIDO, F. J.; STINCO, C. M.; TERRAB, A.; DÍAZ-MORENO, C.; FUENMAYOR, C.; HEREDIA, F. J. Carotenoid profile determination of bee pollen by advanced digital image analysis. **Computers and Electronics in Agriculture**, v. 175, p. 105601, 2020.

HADIMANI, L.; MITTAL, N. Development of a computer vision system to estimate the colour indices of Kinnow mandarins. **Journal of food science and technology**, v. 56, n. 4, p. 2305-2311, 2019.

HEUZÉ, V.; TRAN, G.; ARCHIMÈDE, H.; RÉGNIER, C.; BASTIANELLI, D.; LEBAS, F. **Cassava roots**. Feedipedia, a programme by INRA, CIRAD, AFZ and FAO, 2016.

IBGE [INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA], 2019. **Levantamento Sistemático da Produção Agrícola**. Disponível em: <https://biblioteca.ibge.gov.br/visualizacao/periodicos/2415/epag_2019_dez.pdf>. Acesso em: 02 de Junho de 2021.

IITA. **Cassava in tropical Africa**: a reference manual. Ibadan, 1990, 184p.

IKEOGU, U. N.; DAVRIEUX, F.; DUFOUR, D.; CEBALLOS, H.; EGESI, C. N.; JANNINK, J. L. Rapid analyses of dry matter content and carotenoids in fresh cassava roots using a portable visible and near infrared spectrometer (Vis/NIRS). **PLoS one**, v. 12, n. 12, p. 1-17, 2017.

JARVIS, A.; RAMIREZ-VILLEGAS, J.; CAMPO, B. V. H.; NAVARRO-RACINES, C. Is cassava the answer to African climate change adaptation?. **Tropical Plant Biology**, v. 5, n. 1, p. 9-29, 2012.

JIMÉNEZ, A.; FABRA, M. J.; TALENS, P.; CHIRALT, A. Edible and biodegradable starch films: a review. **Food and Bioprocess Technology**, v. 5, n. 6, p. 2058-2076, 2012.

KHIEU, B.; CHHAY, T.; OGLE, R. B.; PRESTON, T. R. **Research on the use of cassava leaves for livestock feeding in Cambodia**. Proceeding of the regional workshop on “The Use of Cassava Roots and Leaves for On- Farm Animal Feeding”, Hue, Vietnam, p. 17-19, 2005.

KHOSHROO, A.; AREFI, A.; KHODAEI, J. Detection of red tomato on plants using image processing techniques. **Agricultural Communications**, v. 2, n. 4, p. 9-15, 2014.

KOUNDINYA, A. V. V.; HEGDE, V.; SHEELA, M. N.; CHANDRA, C. V. Evaluation of Cassava Varieties for Tolerance to Water Deficit Stress conditions. **JOURNAL OF ROOT CROPS**, v. 44, n. 1, p. 70-75, 2018.

KURTULMUS, F.; LEE, W. S.; VARDAR, A. Immature peach detection in colour images acquired in natural illumination conditions using statistical classifiers and neural network. **Precision agriculture**, v. 15, n. 1, p. 57-79, 2014.

LATIF, S.; MÜLLER, J. Potential of cassava leaves in human nutrition: a review. **Trends in Food Science & Technology**, v. 44, n. 2, p. 147-158, 2015.

LEMOS, P. V. F.; BARBOSA, L. S.; RAMOS, I. G.; COELHO, R. E.; DRUZIAN, J. I. Characterization of amylose and amylopectin fractions separated from potato, banana, corn, and cassava starches. **International journal of biological macromolecules**, v. 132, p. 32-42, 2019.

LÉOTARD, G.; DUPUTIÉ, A.; KJELLBERG, F.; DOUZERY, E. J.; DEBAIN, C.; DE GRANVILLE, J. J.; MCKEY, D. Phylogeography and the origin of cassava: new insights from the northern rim of the Amazonian basin. **Molecular Phylogenetics and Evolution**, v. 53, n. 1, p. 329-334, 2009.

LI, M.; FRANK, M. H.; MIGICOVSKY, Z. ColourQuant: a high-throughput technique to extract and quantify colour phenotypes from plant images. **arXiv preprint arXiv:1903.01652**, 2019.

LI, M.; ZI, X.; ZHOU, H; LV, R.; TANG, J.; CAI, Y. Silage fermentation and ruminal degradation of cassava foliage prepared with microbial additive. **AMB Express** **9**, v. 9, n. 1, p. 1-6, 2019.

LU, Y.; DING, Y.; WU, Q. Simultaneous saccharification of cassava starch and fermentation of algae for biodiesel production. **Journal of Applied Phycology**, v. 23, n. 1, p. 115-121, 2011.

MALDONADO-ROBLEDO, G.; RODRIGUEZ-BUSTAMANTE, E.; SANCHEZ-CONTRERAS, A.; RODRIGUEZ-SONOJA, R.; SANCHEZ, S. Production of tobacco aroma from lutein. Specific role of the microorganisms involved in the process. **Applied Microbiology and Biotechnology**, v. 62, n. 5-6, p. 484-488, 2003.

MAZIYA-DIXON, B.; DIXON, A. G.O.; SSEMAKULA, G. Changes in total carotenoid content at different stages of traditional processing of yellow-fleshed cassava genotypes. **International journal of food science & technology**, v. 44, n. 12, p. 2350-2357, 2009.

NHASSICO, D.; MUQUINGUE, H.; CLIFF, J.; CUMBANA, A.; BRADBURY, J. H. Rising African cassava production, diseases due to high cyanide intake and control measures. **Journal of the Science of Food and Agriculture**, v. 88, n. 12, p. 2043-2049, 2008.

OLIVEIRA, E. J. de; FERREIRA, C. F.; SANTOS, V. da S.; JESUS, O. N. de; OLIVEIRA, G. A. F.; SILVA, M. S. da Potential of SNP markers for the characterization of Brazilian cassava germplasm. **Theoretical and Applied Genetics**, v. 127, n. 6, p. 1423-1440, 2014.

OLIVEIRA, E. J.; OLIVEIRA FILHO, O. S.; SANTOS, V. S. Classification of cassava genotypes based on qualitative and quantitative data. **Genetics and Molecular Research**, v. 14, n. 1, p. 906-924, 2015.

OLSEN, K. M.; SCHAAL, B. A. Evidence on the origin of cassava: phylogeography of *Manihot esculenta*. **Proceedings of the National Academy of Sciences**, v. 96, n. 10, p. 5586-5591, 1999.

OLURANTI, O. M.; BADEJO, A. A.; FAGBEMI, T. N. Processing effects on the total carotenoid content and acceptability of food products from cultivars of biofortified cassava (*Manihot esculenta* Crantz). **Applied Tropical Agriculture**, v. 20, n. 2, p.104-109, 2016.

ORORORO, O. C.; TONUAKARI, N. J.; AVWIOROKO, O. J.; EZEDOM, T. Effect of supplementation of animal feed with dried cassava (*Manihot esculenta*) peels, and stems of *Vernonia amygdalina* and *Pennisetum purpureum* on some biochemical parameters in pigs. **NISEB Journal**, v. 14, n. 4, p. 177–183, 2014.

OSPINA, B.; CEBALLOS, H. **Cassava in the third millennium: modern production, processing, use, and marketing systems**. CIAT, Cali, Colombia. 574 p., 2012.

PACE, B. CEFOLA, M.; RENNA, F.; RENNA, M.; SERIO, F.; ATTOLICO, G. Multiple regression models and computer vision systems to predict antioxidant activity and total phenols in pigmented carrots. **Journal of food engineering**, v. 117, n. 1, p. 74-81, 2013.

PARMAR, A.; STURM, B.; HENSEL, O. Crops that feed the world: Production and improvement of cassava for food, feed, and industrial uses. **Food Security**, v. 9, n. 5, p. 907-927, 2017.

RABBI, I. Y.; UDOH, L. I.; WOLFE, M.; PARKES, E. Y.; GEDIL, M. A.; DIXON, A.; RAMU, P.; JANNINK, J-L.; KULAKOW, P. Genome-wide association

mapping of correlated traits in cassava: dry matter and total carotenoid Content. **The plant genome**, v. 10, n. 3, 2017.

RAHMAWATI, R. S.; SUKMA, D.; ARDIE, S. W.; SUDARSONO, S. Postharvest physiological deterioration in cassava: potential problems, possible inhibition, and resistant level identification. In: **IOP Conference Series: Earth and Environmental Science**. IOP Publishing, p. 012035, 2021.

RAMALHO, M. A. P.; SANTOS, J. B. dos; PINTO, C. A. B. P. **Genética na agropecuária**. 4^a ed. Editora UFLA, Lavras, 2008. 463p.

RIJSSEN, F. W. J. V.; MORRIS, E. J.; ELOFF, J. N. Food safety: importance of composition for assessing genetically modified cassava (*Manihot esculenta* Crantz). **Journal of agricultural and food chemistry**, v. 61, n. 35, p. 8333-8339, 2013.

RODRIGUEZ-AMAYA, D. B.; KIMURA, M.; AMAYA-FARFAN, J. **Fontes brasileiras de carotenoides**. Brasília: Ministério de Meio Ambiente, 2008. p. 100.

SALTZMAN, A.; BIROL, E.; OPARINDE, A.; ANDERSSON, M. S.; ASARE-MARFO, D.; DIRESSIE, M. T.; GONZALEZ, C.; LIVIDINI, K.; MOURSI, M.; ZELLER, M. (Availability, production, and consumption of crops biofortified by plant breeding: current evidence and future potential. **Annals of the New York Academy of Sciences**, v. 1390, n. 1, p. 104-114, 2017.

SANAEIFAR, A.; JAFARI, A.; GOLMAKANI, M-T. Fusion of dielectric spectroscopy and computer vision for quality characterization of olive oil during storage. **Computers and Electronics in Agriculture**, v. 145, p. 142-152, 2018.

SÁNCHEZ, T.; CEBALLOS, H.; DUFOUR, D.; ORTIZ, D.; MORANTE, N.; CALLE, F.; FELDE, T. Z.; DOMÍNGUEZ, M.; DAVRIEUX, F. Prediction of

carotenoids, cyanide and dry matter contents in fresh cassava root using NIRS and Hunter color techniques. **Food Chemistry**, v. 151, p. 444–451, 2014.

SÁNCHEZ, T.; MAFLA, G.; MORANTE, N.; CEBALLOS, H.; DUFOUR, D.; CALLE, F.; MORENO, X.; PÉREZ, J.C.; DEBOUCK, D. Screening of starch quality traits in cassava (*Manihot esculenta* Crantz). **Starch-Starke**, v. 61, p. 12-19, 2009.

SAYRE, R.; BEECHING, J. R.; CAHOON, E. B.; EGESI, C.; FAUQUET, C.; FELLMAN, J.; FREGENE, M.; GRUISSEM, W.; MALLOWA, S.; MANARY, M.; MAZIYA-DIXON, B.; MBANASO, A.; SCHACHTMAN, D. P.; SIRITUNGA, D.; TAYLOR, N.; VANDERSCHUREN, H.; ZHANG, P. The BioCassava plus program: biofortification of cassava for sub-Saharan Africa. **Annual review of plant biology**, v. 62, p. 251-272, 2011.

SELVARAJ, M. G.; VALDERRAMA, M.; GUZMAN, D.; VALENCIA, M.; RUIZ, H.; ACHARJEE, A. Machine learning for high-throughput field phenotyping and image processing provides insight into the association of above and below-ground traits in cassava (*Manihot esculenta* Crantz). **Plant methods**, v. 16, n. 1, p. 1-19, 2020.

SONATI, J.G.; VILARTA, R.; SILVA, C.C. Influências culinárias e diversidade cultural da identidade brasileira: Imigração, Regionalização e suas Comidas. **Qualidade de Vida e Cultura Alimentar**. 1ª ed. v.1, Campinas: IPES, 2009, p.137-147.

STINCO, C. M.; RODRÍGUEZ-PULIDO, F. J.; ESCUDERO-GILETE, M. L.; GORDILLO, B.; VICARIO, I. M.; MELÉNDEZ-MARTÍNEZ, A. J. Lycopene isomers in fresh and processed tomato products: Correlations with instrumental color measurements by digital image analysis and spectroradiometry. **Food research international**, v. 50, n. 1, p. 111-120, 2013.

STROBEL, M.; TINZ, J.; BIESALSKI, H-K. The importance of β -carotene as a source of vitamin A with special regard to pregnant and breastfeeding women. **European journal of nutrition**, v. 46, n. 9, p. 1-20, 2007.

WALTER, J.; EDWARDS, J.; CAI, J.; MCDONALD, G.; MIKLAVCIC, S. J.; KUCHEL, H. High-throughput field imaging and basic image analysis in a wheat breeding programme. **Frontiers in plant science**, v. 10, p. 449, 2019.

WASONGA, D. O.; YAW, A.; KLEEMOLA, J.; ALAKUKKU, L.; MÄKELÄ, P. S. Red-green-blue and multispectral imaging as potential tools for estimating growth and nutritional performance of cassava under deficit irrigation and potassium fertigation. **Remote Sensing**, v. 13, n. 4, p. 598, 2021.

YE, X.; IZAWA, T.; ZHANG, S. Rapid determination of lycopene content and fruit grading in tomatoes using a smart device camera. **Cogent Engineering**, v. 5, n. 1, p. 1504499, 2018.

ZHU, T.; JACKSON, D. S.; WEHLING, R. L.; GEERA, B. Comparison of amylose determination methods and the development of a dual wavelength iodine binding technique. **Cereal Chemistry**, v. 85, n. 1, p. 51-58, 2008.

ARTIGO 1**DIVERSIDADE FENOTÍPICA E SELEÇÃO NO GERMOPLASMA
BIOFORTIFICADO DE MANDIOCA PARA CARACTERÍSTICAS
PRODUTIVAS E DE QUALIDADE DA RAIZ**

DIVERSIDADE FENOTÍPICA E SELEÇÃO NO GERMOPLASMA BIOFORTIFICADO DE MANDIOCA PARA CARACTERÍSTICAS PRODUTIVAS E DE QUALIDADE DA RAIZ

RESUMO: O aumento do teor de carotenoides (TCC) e a melhoria de outras características da qualidade das raízes tem sido o foco da biofortificação de mandioca. Este estudo objetivou: i) avaliar a variabilidade genética para TCC, bem como para atributos produtivos e de qualidade das raízes no germoplasma de mandioca, ii) estimar correlações potencialmente úteis para seleção, e iii) selecionar parentais e estimar ganho genéticos potenciais. Foram utilizados dados de 2011 a 2020 de 265 genótipos de mandioca com raízes de coloração creme e amarela para características como teor de matéria seca pelos métodos gravimétrico (*DMC.Grav*, %) e secagem em estufa (*DMC.OD*, %), produtividade da parte aérea (*PPA*, t.ha⁻¹), raízes frescas (*PTR*, t.ha⁻¹) e secas (*PROD.Dry*, t.ha⁻¹), índice de colheita (*IC*, %), número médio de raízes por planta (*NRP*), teor de amido (*AMD*, %), cor da polpa das raízes (*CorPolpa*), teor de compostos cianogênicos (*HCN*) e de carotenoides totais (TCC, µg.g⁻¹). A distribuição dos BLUPs demonstrou grande variação fenotípica para todas as características de qualidade da raiz e agronômicas no germoplasma de polpa amarela. Seis grupos bastante distintos em relação às características produtivas de qualidade das raízes, sobretudo TCC, DMC e PTR foram identificados. Apenas TCC apresentou alta herdabilidade ($h^2=0,72$), enquanto as demais características tiveram baixa a média magnitude (h^2 entre 0,21 e 0,60). O TCC foi fortemente correlacionado com cor de polpa ($r=0,70$), mas praticamente nula e sem significância para DMC. A análise *network* identificou uma clara separação entre os atributos agronômicos e de qualidade das raízes de mandioca, com algumas interconexões entre elas. A seleção dos 30 genótipos para recombinação no programa de melhoramento de mandioca tem potencial para elevar o TCC em 37% e reduzir o HCN em -19,4%, além de aumentar a PTR e PROD.Dry em 37,7% e 40%, respectivamente. Este é o primeiro estudo consolidado sobre o potencial do germoplasma para o desenvolvimento de variedades biofortificadas de mandioca no Brasil.

Palavras-chave: Carotenoides, melhoramento genético, qualidade de raiz.

PHENOTYPICAL DIVERSITY AND SELECTION IN BIOFORTIFIED CASSAVA GERMPLASM FOR YIELD AND QUALITY ROOT TRAITS

ABSTRACT: Increasing carotenoid content (TCC) and improving other root quality has been the focus of cassava biofortification. This study aimed to: i) evaluate the genetic variability for TCC, as well as for root yield and root quality attributes in cassava germplasm, ii) estimate potentially useful correlations for selection, and iii) select parents for breeding and estimate the potential genetic gain. Historical data from 2011 to 2020 of 265 cassava genotypes with cream and yellow roots were analyzed for characteristics such as dry matter content by gravimetric methods (DMC.Grav, %) and oven dry (DMC.OD, %), shoot yield (PPA, t.ha⁻¹), fresh root yield (PTR, t.ha⁻¹) and dry root yield (PROD.Dry, t.ha⁻¹), harvest index (IC, %), average number of roots per plant (NRP), starch content (AMD, %), root pulp color (CorPolpa), cyanogenic compounds (HCN) and total carotenoids content (TCC, µg.g⁻¹). The distribution of BLUPs showed great phenotypic variation for all root quality and agronomic traits in yellow pulp germplasm. Six very distinct groups were formed in relation to the productive characteristics of root quality, mainly TCC, DMC and PTR. Only TCC showed high broad-sense heritability ($h^2=0.72$), while the other traits had low to medium magnitude ($h^2=$ between 0.21 and 0.60). TCC was strongly correlated with pulp color ($r=0.70$), but practically null and without significance for dry matter content in the roots. The network analysis identified a clear separation between the agronomic and quality attributes of cassava roots, with some interconnections between them. The selection of the 30 genotypes for recombination in the cassava breeding program has the potential to raise TCC by 37% and reduce the content of cyanogenic compounds by -19.4%, in addition to increasing PTR and PROD.Dry by 37.7 % and 40%, respectively. This is the first consolidated study on the potential of germplasm for the development of biofortified cassava varieties in Brazil.

Keywords: Carotenoids, genetic improvement, root quality.

1. INTRODUÇÃO

A mandioca (*Manihot esculenta* Crantz) é uma cultura de raiz amilácea amplamente cultivada no Sudeste Asiático, América Latina, Caribe e África Subsaariana para consumo humano e animal e como matéria-prima para biocombustíveis e diversos outros produtos industrializados com ampla gama de usos (HOWELER; LUTALADIO; THOMAS, 2013). A Nigéria é líder em produção de raízes de mandioca, seguida pela República Democrática do Congo, Tailândia e Gana, com uma produção de 59,19; 40,05; 31,07; e 22,44 milhões de toneladas, respectivamente, de raiz em 2019 (FAO, 2020). O Brasil se destaca como o 5º maior produtor mundial, com produção de 17,49 milhões de toneladas, colhida em cerca de 1,19 milhões de hectares, com um rendimento médio nacional de cerca de 14,70 t.ha⁻¹ (FAO, 2020). Embora, haja grandes variações na produtividade nacional de mandioca, pois a média da região Nordeste é de 9,80 t.ha⁻¹, enquanto a região Sul do Brasil está em 21,71 t.ha⁻¹ (IBGE, 2019).

As raízes da mandioca são uma excelente fonte de calorias devido ao acúmulo de amido (MONTAGNAC; DAVIS; TANUMIHARDJO, 2009), sendo este o principal produto explorado comercialmente na mandioca, cujos usos na indústria alimentícia, papel e celulose, têxtil, petróleo, dentre outras é bastante reconhecido (BALAGOPALAN, 2002; FAO, 2018). A mandioca é a segunda fonte mais importante de amido no mundo (STAPLETON, 2012). Além do amido, a raiz da mandioca possui água, fibras, frações de proteína, gordura, elementos minerais (Zn, Fe, N, Ca, P, K e Mg) e vitaminas (B1, B2, B3, C e β -caroteno) (SAYRE et al., 2011; PARMAR; STURM; HENSEL, 2017; BAYATA; GURE; KEDIR, 2018).

A mandioca possui alta resiliência ambiental e é capaz de garantir níveis mínimos de produtividade mesmo quando submetida a diversos estresses ambientais, a exemplo do hídrico, cultivo em solo de baixa fertilidade natural e a manejos com baixo uso de tecnologias produtivas (BURNS et al., 2010; JARVIS et al., 2012). É uma cultura de fácil adaptação aos cenários associados às futuras mudanças climáticas, tendo assim um alto potencial para ser uma cultura de segurança alimentar em regiões menos favorecidas

economicamente (FAO, 2018). Os programas de melhoramento da mandioca buscam identificar e selecionar genótipos com características desejáveis a exemplo da produtividade de raízes e amido, além da resistência a pragas e doenças (OLIVEIRA; OLIVEIRA FILHO; SANTOS, 2015), menor toxicidade (PARMAR; STURM; HENSEL, 2017), tolerância ao déficit hídrico (KOUNDINYA et al., 2018) e melhor conteúdo nutricional (ESUMA et al., 2012; CEBALLOS et al., 2013).

A biofortificação de culturas básicas é uma iniciativa mundial cujo objetivo é garantir a segurança alimentar da população, fornecendo os nutrientes essenciais para o desenvolvimento humano por meio da alimentação básica (MAYER; PFEIFFER; BEYER, 2008; SALTZMAN et al., 2017; BOUIS; SALTZMAN, 2017; ALAMU et al., 2019). O processo da biofortificação em mandioca visa aumentar o conteúdo de carotenoides por meio de técnicas convencionais de melhoramento, tirando proveito da variabilidade genética natural da espécie (ESUMA et al., 2012; CEBALLOS et al., 2013), assim como pela transformação genética (WELSCH et al., 2010; FAILLA et al., 2012).

Os carotenoides desempenham papel fundamental para saúde humana (CUEVAS et al., 2010), pois o β -caroteno é responsável por atuar como precursor da vitamina A, possui propriedades antioxidantes (SHETE; QUADRO, 2013) e é o carotenoide mais abundante na mandioca (CEBALLOS et al., 2017). Portanto, a mandioca é uma espécie candidata a atuar como fonte de carotenoides, sobretudo em regiões onde seu consumo é alto devido a questões culturais tradicionais e também pela dificuldade de se cultivar outras espécies, a exemplo da região semiárida no Nordeste do Brasil.

A maioria das variedades lançadas no Brasil até 2002 tinha coloração de polpa branca (FUKUDA et al., 2002). Nesse período, o teor de carotenoides em clones melhorados e *landraces* no germoplasma de mandioca mantido pelo CIAT, incluindo clones oriundos do Brasil, apresentava uma variação de 1,02 a 10,40 $\mu\text{g}\cdot\text{g}^{-1}$ de conteúdo de carotenoides nas raízes frescas (CHAVÉZ et al., 2005). Nos últimos anos, as novas variedades biofortificadas lançadas no Brasil apresentaram teores de carotenoides totais variando entre 3,3 e 8,7 $\mu\text{g}\cdot\text{g}^{-1}$ de β -caroteno (FUKUDA et al., 2005; FUKUDA; PEREIRA, 2005; FUKUDA et al., 2009). Entretanto, recentes estudos têm demonstrado uma variação entre 0,20

a 30,84 $\mu\text{g}\cdot\text{g}^{-1}$ no germoplasma de mandioca (SÁNCHEZ et al., 2014; RABBI et al., 2017). A presença de variabilidade genética para o teor de carotenoides totais tem permitido o desenvolvimento de germoplasma biofortificado, usando estratégias convencionais de melhoramento, para tratar a deficiência de vitamina A na África Subsaariana (SAYRE et al., 2011; NJOKU et al., 2014; TALSMA et al., 2016), e na América Latina (MEENAKSHI et al., 2010; CEBALLOS et al., 2013).

A possibilidade de aliar variedades biofortificadas com características culinárias desejáveis e alto desempenho agrônomico tem demandado muitas pesquisas voltadas para o melhoramento genético da mandioca (MEZZETE et al., 2009; BELALCAZAR et al., 2016; FUHRMANN et al., 2019; PARKES et al., 2020). Entretanto, o programa de biofortificação em mandioca deve focar no desenvolvimento de variedades de mesa, considerando que a retenção de carotenoides em amostras de raízes cozidas está entre 95-100% (TALEON et al., 2019). Neste caso, além do alto teor de carotenoides as variedades de mesa biofortificadas precisam ter baixo teor de compostos cianogênicos, alto teor de matéria seca nas raízes, associadas à alta performance agrônômica.

As variedades de mandioca “brava” ou amarga possuem elevado teor de compostos cianogênicos (linamarina e lotaustralina) em toda a planta, os quais são efetivamente reduzidos com diferentes métodos de processamento e fermentação (PARMAR; STURM; HENSEL, 2017). De acordo com a concentração de glicosídeos cianogênicos nas raízes, os genótipos podem ser classificados em mandioca “brava” (≥ 100 mg compostos cianogênicos kg^{-1}) ou mandioca “mansa” (≤ 100 mg compostos cianogênicos kg^{-1}) (MCKEY et al., 2010).

O teor de matéria seca nas raízes podem variar de 8,4 a 52,5% (SÁNCHEZ et al., 2014; OLIVEIRA et al., 2015; OLIVEIRA, et al., 2017; RABBI et al., 2017), com ampla variação natural para melhorar esta característica. Entretanto, diversos relatos em painéis de germoplasma oriundos da África (clones melhorados e variedades locais), tem demonstrado a existência de correlações negativas (-0,22 a -0,59) entre teor de carotenoides e teor de matéria seca nas raízes (AKINWALE et al., 2010; ESUMA et al., 2012; NJOKU et al., 2015; RABBI et al., 2017). Recentemente, Rabbi et al. (2017), analisaram

um painel de 672 clones africanos de mandioca e identificaram dois loci no cromossomo 1 nas posições 24.1 e 30.5 Mbp, associados à cor das raízes de mandioca (altamente correlacionada com teor de carotenoides) e um único locus para o teor de matéria seca na região próxima a 24,1 Mbp no cromossomo 1. Estes autores relataram, portanto, a ocorrência de ligação física entre estas duas características no germoplasma de mandioca africano.

A garantia da adoção de variedades biofortificadas de mandioca depende da geração de progênies que permitam aumento simultâneo do teor de carotenoides e características agrônômicas que são essências para garantir a comercialização das raízes (alto teor de matéria seca, baixo teor de compostos cianogênicos, raízes com tamanho e formato adequados) e ao mesmo tempo a rentabilidade dos agricultores (alta produtividade de raízes, resistência pragas e doenças e adequação a sistemas de plantio minimamente mecanizados).

Um dos principais passos para a inclusão de clones como parentais em blocos de cruzamentos é a caracterização do seu potencial produtivo, de resistência a doença, adequação ao manejo cultural alvo, e sobretudo quanto a qualidade das raízes. Esta etapa permite identificar combinações parentais que podem maximizar as recombinações para gerar progênies segregantes com alta variabilidade genética e ao mesmo tempo permitir a introgressão de genes desejáveis (OLIVEIRA et al., 2014). Entretanto, poucos estudos tem se dedicado a uma ampla caracterização do germoplasma de mandioca brasileiro para todas estas características associadas à aceitação de variedades biofortificadas. Portanto, este estudo teve como objetivo avaliar a variabilidade fenotípica para características associadas à qualidade de raiz e produtividade de genótipos biofortificados de mandioca e selecionar os mais promissores para uso no melhoramento da espécie. Também foram analisadas as correlações entre as diversas características avaliadas e o ganho genético potencial com a seleção dos melhores clones para blocos de cruzamentos.

2. MATERIAL E MÉTODOS

2.1 Ensaios de campo para coleta dos dados

Neste estudo foram utilizados dados do programa de melhoramento genético de mandioca da Embrapa Mandioca e Fruticultura, coletados entre os anos de 2011 a 2020. No total foram conduzidos 21 ensaios nas cidades baianas de Cruz das Almas, Laje e Valença entre os anos de 2011 a 2020. Informações detalhadas sobre cada tipo de solo, coordenadas dos ensaios e anos de avaliação estão disponíveis na Tabela 1. De acordo com a classificação de Köppen, o clima das regiões de avaliação dos ensaios é do tipo Af, i.e., clima tropical quente e úmido, com precipitação média anual variando de 1200 a 1500 mm, com maior incidência de chuvas no período compreendido entre março e julho, com temperatura média anual de 24,5°C e umidade relativa do ar de aproximadamente 80%.

Os ensaios de avaliação do germoplasma de mandioca foram instalados em delineamento de blocos aumentados, com parcelas compostas por duas linhas com oito a dez plantas cada (total de 16 a 20 plantas por parcela), espaçadas de 0,90 m entre linhas e 0,80 m entre plantas. Entre 4 e 34 variedades melhoradas da Embrapa e variedades locais foram plantadas como testemunhas comuns nos diferentes blocos aumentados (que variaram de 5 a 22).

O sistema de cultivo convencionalmente utilizado na região foi utilizado para preparar com solo (uma aração, duas gradagens e abertura dos sulcos de plantio com auxílio de uma plantadeira de mandioca, sem discos de cobertura do solo). O plantio de manivas com aproximadamente 16-18 cm de comprimento foi realizado manualmente.

Todos os ensaios foram conduzidos em condições de sequeiro (sem irrigação complementar), seguindo o manejo cultural da região, de acordo com as recomendações de Souza; Souza; Gomes (2006). O plantio foi realizado durante o período chuvoso na região (maio a agosto) de modo a garantir a quantidade mínima de umidade no solo necessária para germinação e estabelecimento da cultura.

Tabela 1. Localização e caracterização dos locais de implementação dos ensaios de avaliação do germoplasma de mandioca.

Código	Ano	Local	Cidade Baianas	Altitude	Coordenadas	Tipo de solo*
AgroV1	2011	Embrapa	Cruz das Almas	215	12°40'36.7"S 39°05'08.0"W	LatAm
AgroV2	2011	Embrapa	Cruz das Almas	210	12°40'32.6"S 39°05'13.7"W	LatAm
BAG-C	2012	Embrapa	Cruz das Almas	200	12°40'47.4"S 39°05'00.2"W	LatAm
AgroV1	2013	Bahiamido	Laje	180	13°06'38.4"S 39°16'20.4"W	LatVemAm
AgroV1	2013	UFRB	Cruz das Almas	210	12°39'25.9"S 39°04'58.8"W	LatAm
BAG-C	2013	Embrapa	Cruz das Almas	217	12°40'22.8"S 39°05'06.1"W	LatAm
BAG-1.2	2014	Embrapa	Cruz das Almas	216	12°40'22.8"S 39°05'01.5"W	LatAm
BAG-1.1	2014	Embrapa	Cruz das Almas	216	12°40'22.8"S 39°05'01.5"W	LatAm
BAG-1	2014	Bahiamido	Laje	180	13°06'39.6"S 39°16'17.6"W	LatAm
BAG-2	2014	Bahiamido	Laje	175	13°06'35.6"S 39°16'19.3"W	LatAm
BAG-3	2014	Bahiamido	Laje	180	13°06'39.6"S 39°16'17.6"W	LatAm
BAG-1.1	2015	Embrapa	Cruz das Almas	216	12°40'22.8"S 39°05'01.5"W	LatAm
BAG-1	2015	Bahiamido	Laje	196	13°06'38.5"S 39°16'49.0"W	LatAm
BAG-2	2015	Bahiamido	Valença	40	13°15'33.5"S 39°14'12.8"W	ArgVemAm
BAG-2	2015	UFRB	Cruz das Almas	210	12°39'16.4"S 39°04'53.4"W	LatAm
BAG-3	2015	Bahiamido	Laje	196	13°06'38.5"S 39°16'49.0"W	LatAm
BAG-4	2015	Embrapa	Cruz das Almas	216	12°40'19.5"S 39°05'02.5"W	LatAm
BAG-1.1	2018	UFRB	Cruz das Almas	223	12°39'51.4"S 39°04'15.7"W	LatAm
BAG-1.1	2019	UFRB	Cruz das Almas	223	12°39'43.5"S 39°04'12.0"W	LatAm
BAG-1.1	2020	UFRB	Cruz das Almas	225	12°39'49.2"S 39°03'58.1"W	LatAm
BAG-1.2	2020	UFRB	Cruz das Almas	225	12°39'49.2"S 39°03'58.1"W	LatAm

*LatAm = Latossolo Amarelo, LatVemAm = Latossolo vermelho amarelo, ArgVemAm = Argissolo vermelho amarelo

2.2 Painel de germoplasma

O painel de acessos de mandioca analisados foi constituído basicamente por genótipos que possuem raízes de coloração creme a amarela, considerando que a cor amarela das raízes de mandioca está diretamente associada com o teor de carotenoides (ESUMA et al., 2016). Portanto, foi selecionado um painel com 265 acessos de mandioca constituídos por variedades melhoradas e locais (não melhoradas) (Tabela Suplementar 1). A maioria deste germoplasma tem sua origem na Colômbia (6%) e Brasil (3%, 34%, 37%, 4%, 6% nas regiões Centro-Oeste, Norte, Nordeste, Sul e Sudeste, respectivamente, além de outros 9% com origem desconhecida).

2.3 Coleta dos dados agronômicos

A colheita dos ensaios foi realizada manualmente entre 10-12 meses após o plantio e nas avaliações de 2011 a 2020 foram avaliadas as seguintes características:

1) teor de matéria seca nas raízes com base no método gravimétrico (*DMC.Grav*) em %: aproximadamente 5kg de raízes foram limpas para retirar o excesso de solo e cortadas as extremidades. Em seguida foram obtidos o peso das raízes no ar e na água e calculado o *DMC.Grav* com base na fórmula:

$$DM = 158.3 \times \frac{\text{Peso na água}}{\text{Peso no ar} - \text{Peso na água}} - 142, \text{ de acordo com Kawano et al. (1987),}$$

2) teor de matéria seca nas raízes com base na secagem das amostras (*DMC.OD*) em %: entre 3 a 5 raízes de diferentes plantas foram selecionadas e limpas para retirar o excesso de solo. Em seguida, aproximadamente 200 g de raízes oriundas de diferentes posições das raízes foram trituradas para facilitar a secagem em forno a 90° C até peso constante (geralmente em 48 horas). A *DMC.OD* foi determinada pela fórmula: $MS (\%) = 100 - \text{Umidade} (\%)$,

3) produtividade da parte aérea (PPA): obtida pela colheita e remoção das raízes das plantas, pesando somente a parte aérea de todas as plantas na parcela, sendo ajustada para $t.ha^{-1}$,

4) produtividade de raízes frescas (PTR), obtido por meio da pesagem de todas as raízes da parcela, com auxílio de uma balança digital, posteriormente convertido para $t.ha^{-1}$,

5) índice de colheita (IC) em %: relação entre o peso das raízes e da parte aérea, de acordo com a formula: $IC = \frac{PTR}{PTR+PPA} \times 100$,

6) produtividade de raízes secas (*PROD.Dry*) em $t.ha^{-1}$: multiplicação de PTR \times *DMC.Grav*,

7) número médio de raízes por planta (NRP): obtido pela contagem de todas as raízes da parcela dividido pelo número de plantas colhidas,

8) teor de amido nas raízes (AMD) em %: o amido de aproximadamente 1 kg de raízes oriundas de diferentes plantas da parcela foi extraído de acordo com Vasconcelos et al. (2017),

9) cor da polpa das raízes (CorPolpa): classificadas em 1 = raízes brancas, 2 = raízes de cor creme e 3 = raízes de cor amarela,

10) teor de compostos cianogênicos (HCN): obtido pelo método do picrato (FUKUDA et al., 2010).

2.4 Análise do teor de carotenoides totais

Para análise do teor de carotenoides totais (TCC), *DMC.OD* e AMD as amostras foram conduzidas até o Laboratório de Práticas Culturais de Mandioca, nas primeiras horas da manhã. Todas as raízes foram lavadas, descascadas e cotadas em pequenos pedaços antes de iniciar as análises. As avaliações para TCC foram realizadas apenas nos anos de 2019 e 2020.

Para TCC, as raízes foram manipuladas sempre em condições de baixa luminosidade. Foram coletadas duas amostras contendo 10g, 15g ou 25g de raiz triturada (a depender da intensidade da coloração da polpa) e uma amostra reserva de 60 g, em potes de vidro com tampas, envolvidos com papel alumínio. O TCC foi quantificado conforme descrito no HarvestPlus Handbook for Carotenid Analysis (RODRIGUEZ-AMAYA; KIMURA, 2004). Primeiramente foram extraídos os pigmentos da amostra de mandioca triturada descongelada com aproximadamente 50 mL de acetona, por meio da moagem utilizando o homogeneizador Ultra Turrax.

A massa contendo a amostra e acetona foi filtrada através de um funil de Buchner com auxílio de uma bomba de vácuo, e o resíduo retido no funil foi lavado com acetona até não apresentar nenhuma coloração. O extrato contendo apenas o pigmento e acetona, reservado no balão de sucção, foi transferido para um funil de separação contendo éter de petróleo (a quantidade variável dependendo da coloração da amostra), e adicionada solução salina lentamente (cerca de 250 mL) para que houvesse uma separação das fases. A fase aquosa foi descartada e o procedimento de lavagem com solução salina foi repetido por cinco vezes até que restasse apenas éter de petróleo e pigmento. Esse novo extrato foi filtrado em funil com lã de vidro e sulfato de sódio anidro para um balão volumétrico âmbar e completado seu volume com éter de petróleo. Uma alíquota de cada amostra foi utilizada para determinação do TCC com uso de espectrofotômetro (UV-VIS *Thermo Scientific*, modelo Genesis 10S) ajustado para absorvância a 450 nm.

O TCC foi calculado usando a seguinte fórmula: $TCC(ug.g^{-1}) = \frac{A \times V(ml) \times 10^4}{A_{1cm}^{1\%} \times P(g)}$, em que A é a absorvância; $V(ml)$ é o volume total do extrato em mililitros; $P(g)$ é o peso da amostra em gramas; e $A_{1cm}^{1\%}$ é igual ao coeficiente de extinção do β -caroteno em éter de petróleo (2592).

2.5 Análise dos dados

Utilizou-se abordagem de modelos mistos para análise dos dados de acordo com a formula: $Y_{ijk} = \mu + E_i + B_{(i)j} + G_k + GE_{ik} + \varepsilon_{ijk}$, em que Y_{ijk} é o valor observado do genótipo k no bloco j , do ambiente i ; μ é a constante geral do experimento; E_i é o efeito do ambiente i (aleatório), $i = 1, 2, \dots, n$; $B_{(i)j}$ é o efeito do bloco j , $j = 1, 2, \dots, n$, dentro do ambiente i ; $G_k = T_{k'} + T_{(j)k}$, sendo G_k o efeito do genótipo, $T_{k'}$ o efeito fixo do tratamento comum, $k' = 1, 2, 3, \dots$, e $T_{(j)k}$ o efeito aleatório do tratamento regular k dentro do bloco j . $T_{(i)k} \sim NID(0, \sigma_{T(j)k}^2)$; GE_{ik} = efeito da interação do ambiente i com o genótipo k (aleatório). $GE_{ik} \sim NID(0, \sigma_{GA(j)k}^2)$; ε_{ijk} é o efeito aleatório do erro experimental $\sim NID(0, \sigma^2)$.

Os efeitos do modelo foram estimados com o pacote *lme4* (BATES et al., 2015) do software R versão 4.03 (R CORE TEAM, 2021). Os componentes de variância foram estimados pelo método REML (*restricted maximum*

likelihood) e em seguida obtidos os BLUPs (*best linear unbiased predictor*) para os efeitos aleatórios. A significância de cada efeito do modelo foi testada com base na análise de deviance, de acordo com o teste de razão de verossimilhança (LRT), usando a distribuição de χ^2 a 1% de probabilidade, com um grau de liberdade.

A herdabilidade no sentido amplo foi estimada por $h^2 = \frac{\sigma_g^2}{\sigma_g^2 + \sigma_{ga}^2 + \sigma_e^2}$, enquanto que a herdabilidade no sentido amplo por parcela (*clonal mean heritability*) na média de genótipos foi obtida pela fórmula $h_m^2 = \frac{\sigma_g^2}{\sigma_g^2 + \frac{\sigma_{ga}^2}{a} + \frac{\sigma_e^2}{ra}}$, em que σ_g^2 é a variância do efeito de genótipos, σ_{ga}^2 é a variância da interação genótipo x ensaios, σ_e^2 é a variância do erro, a é o número de ensaios, e r é o produto do número de repetições ajustado pelo número de ensaios.

As correlações entre as características produtivas, qualidade de raízes e teor de carotenoides foram estimadas com base na correlação de Pearson, cuja hipótese de correlação igual a zero foi analisada pelo teste t com $n-2$ graus ($p < 0,05$), utilizando o pacote *corrplot* (WEI; SIMKO, 2021) do software R versão 4.03 (R CORE TEAM, 2021). O correlograma com agrupamento hierárquico foi criado para exibir a magnitude e a direção das correlações. Além disso, a correlação *network* foi realizada utilizando o pacote *qgraph* (EPSKAMP et al., 2012) do software R versão 4.03 (R CORE TEAM, 2021).

O número de agrupamentos do germoplasma de mandioca de polpa amarela com base em diversos atributos associados à qualidade de raízes e aspectos agronômicos foram determinados com base em sucessivos K-means e número crescente de grupos (k variando de 2 a 15) após transformar o conjunto de dados por análise de componentes principais usando a função *find.clusters* do pacote *adeget* (JOMBART; DEVILLARD; BALLOUX, 2010) na versão R 4.0.4. O número de clusters foi comparado usando o Critério de Informação Bayesiano (BIC) de adequação, e a solução de clusterização mais adequada foi representada como aquela com menor BIC. Após determinar o número de grupos mais adequado para representar a diversidade do germoplasma, o dendrograma foi construído com auxílio do pacote *circlize* (GU et al., 2014) na versão R 4.0.4.

Boxplots obtidos com uso do pacote ggstatsplot (PATIL, 2018) na versão R 4.0.4 foram utilizados para visualização da distribuição e existência de diferenças entre os agrupamentos para os atributos agrônômicos e de qualidade de raízes.

Foram selecionados 30 genótipos de mandioca para recombinação e geração de progênes melhoradas de mandioca de mesa biofortificada (alto do teor de carotenoides e baixo teor de compostos cianogênicos), associada à elevada performance agrônômica. Para isso foi utilizado o seguinte índice de seleção baseado na soma de ranks (MULAMBA; MOCK, 1978) e pesos predefinidos:

$$IC = (DMC.Grav \times 5) + (DMC.OD \times 5) + (TCC \times 30) + (PTR \times 10) + (PPA \times 5) + (IC \times 5) + (PROD.Dry \times 10) + (NRP \times 10) + (AMD \times 5) + (CorPolpa \times 30) + (HCN \times -30),$$

que se refere ao somatório dos BLUPs de cada característica multiplicados pelos seus respectivos pesos econômicos.

3 RESULTADOS

3.1 Variação e parâmetros genéticos para teor de carotenoides e características produtivas e de qualidade de raízes

A distribuição dos BLUPs somados ao intercepto (deste ponto em diante, denominado apenas de BLUP) para todas as características é apresentada na Figura 1. A amplitude de variação dos BLUPs para TCC foi bastante elevada (0,075 a 13,08 $\mu\text{g.g}^{-1}$), com média de 4,09 $\mu\text{g.g}^{-1}$. O mesmo foi observado para outras características associadas à qualidade das raízes como AMD (amplitude de 13,29 a 30,93% e média de 25,36%), CorPolpa (amplitude de 1 a 3 no score, com média de 2,11) e HCN (amplitude no score entre 2,46 a 7,97, e média de 6,3). Isso demonstra que a maioria dos genótipos avaliados são de coloração creme e ainda possuem alto teor de compostos cianogênicos. Esta afirmação pode ser evidenciada pela maior distribuição dos genótipos à direita do gráfico de densidade, ou seja, alto HCN. Para as demais características, observa-se uma distribuição próxima à normalidade.

Em relação ao teor de matéria seca nas raízes, a amplitude dos dados obtidos pelo método gravimétrico (amplitude do *DMC.Grav* entre 26,61 a

39,57% e média de 34,22%) foi muito menor do que o método de secagem das amostras (amplitude do *DMC.OD* entre 14,07 a 52,67% e média de 39,99%). Embora o método gravimétrico seja mais simples e rápido na obtenção dos teores de matéria seca, o *DMC.Grav* é uma estimativa obtida de forma indireta com base na regressão proposta por Kawano et al. (1987), e por isso, pode haver erros implícitos a esta metodologia. Por outro lado, o *DMC.OD* é uma análise direta para obtenção dos teores de matéria seca em mandioca.

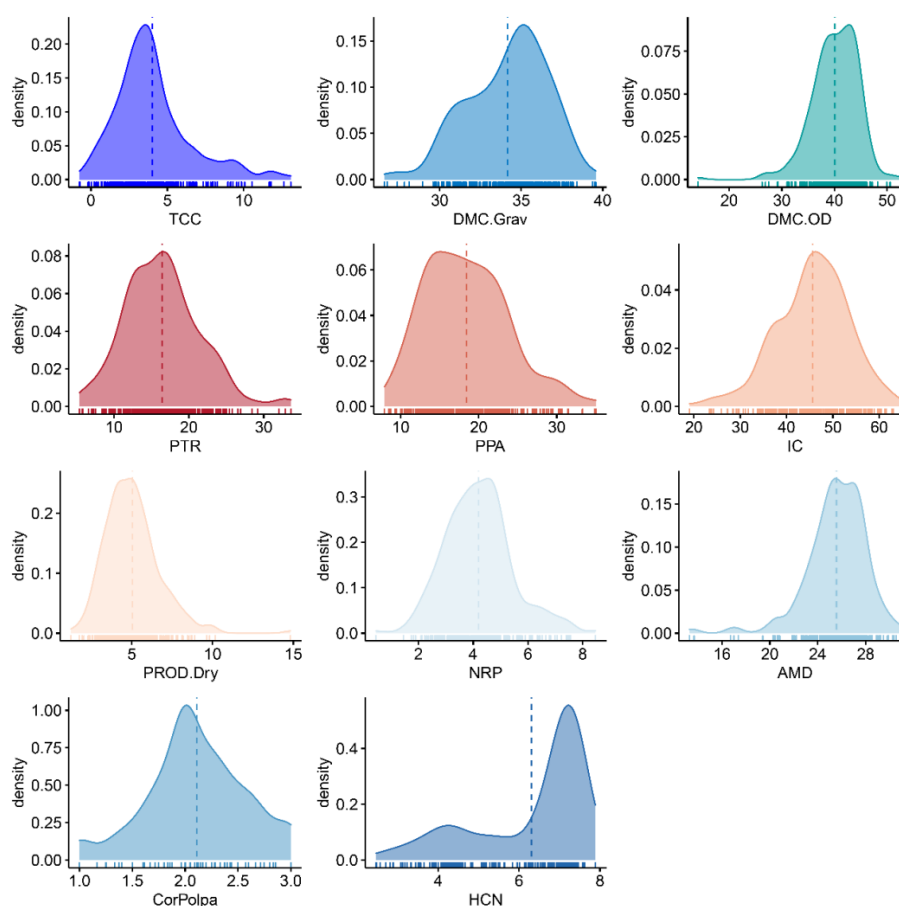


Figura 1. Gráficos de distribuição dos *best linear unbiased prediction* (BLUP) somados à média geral, para diversos atributos associados à qualidade das raízes e produtividade de genótipos de mandioca de raízes amarelas. TCC: teor de carotenoides totais; *DMC.Grav*: teor de matéria seca pelo método gravimétrico; *DMC.OD*: teor de matéria seca pelo método da secagem em estufa; PTR: produtividade de raízes frescas; PPA: produtividade da parte aérea; IC: índice de colheita; *PROD.Dry*: produtividade de raízes secas; NRP: número de raízes; AMD: teor de amido; CorPolpa: cor da polpa das raízes; HCN: teor de compostos cianogênicos.

A variação fenotípica no germoplasma de mandioca de polpa amarela para diversas características agrônômicas, também evidenciaram ampla

variação para PTR (amplitude de 5,41 a 33,62 t.ha⁻¹ e média de 16,49 t.ha⁻¹), PPA (amplitude de 7,92 a 35,00 e média de 18,57 t.ha⁻¹), IC (amplitude de 19,05 a 64,64% e média de 45,40%), *PROD.Dry* (amplitude de 1,24 a 14,81 t.ha⁻¹ e média de 5,06 t.ha⁻¹) e NRP (amplitude de 0,45 a 8,46 e média de 4,15 raízes por planta). Portanto, este painel de germoplasma de mandioca de polpa amarela possui elevado potencial para seleção de genótipos para recombinação de características importantes para o programa de melhoramento de mandioca de mesa, a exemplo do alto TCC e DMC e baixo HCN.

A análise deviance indicou significância do efeito de genótipos para todas estas características (Tabela 2), reforçando os achados prévios relacionados à elevada amplitude para o teor de carotenoides e demais atributos associados à qualidade e produtividade de raízes de mandioca. Além disso, o efeito da interação genótipo × ensaio foi significativo para todas as características, exceto TCC, CorPolpa, *DMC.DO* e AMD.

Tabela 2. Resumo da variação fenotípica, estimativas de componentes de variância e herdabilidade para teor de carotenoides totais (TCC), teor de matéria seca pelo método gravimétrico (*DMC.Grav*) e da secagem em estufa *DMC.DO*, produtividade de raízes frescas (PTR) e secas (*PROD.Dry*), produtividade da parte aérea (PPA), índice de colheita (IC), número de raízes (NRP), teor de amido (AMD), teor de compostos cianogênicos (HCN) e cor da polpa das raízes (CorPolpa).

Características	Média	h^2	h_m^2	σ_g^2	σ_{ge}^2	σ_e^2	χ_g^2	χ_{ge}^2
TCC	4,09	0,72	1,00	5,44	0,00	2,16	100,36*	0,00
DMC.Grav	34,22	0,45	0,92	5,26	3,87	2,59	852,71*	288,62*
DMC.DO	39,99	0,54	0,98	13,15	1,68	9,44	48,74*	0,02
PTR	16,49	0,21	0,83	17,28	28,06	37,56	281,36*	209,28*
PPA	18,57	0,22	0,81	19,12	35,14	32,50	287,30*	305,41*
IC	45,40	0,32	0,89	49,42	46,83	56,40	532,64*	151,22*
PROD.Dry	5,06	0,21	0,80	1,78	3,49	3,22	203,33*	227,16*
NRP	4,15	0,32	0,89	0,81	0,80	0,89	41,30*	31,24*
AMD	25,36	0,48	1,00	5,17	0,00	5,59	15,26*	0,00
HCN	6,80	0,60	0,96	1,60	0,53	0,52	128,22*	150,61*

CorPolpa	2,11	0,42	0,99	0,18	0,00	0,25	31,61*	0,00
----------	------	------	------	------	------	------	--------	------

h^2 = herdabilidade no sentido amplo; h_m^2 = herdabilidade no sentido amplo por parcela, das médias do genótipo; σ_g^2 = variância do genótipo; σ_{ge}^2 = variância da interação genótipo x ensaio; σ_e^2 = variância do erro; χ_g^2 = qui-quadrado do efeito do genótipo; χ_{ge}^2 = qui-quadrado do efeito da interação genótipo x ensaio; * = significativo em $p < 0,01$.

Em relação aos parâmetros genéticos, apenas TCC apresentou alta herdabilidade no sentido amplo ($h^2=0,72$), enquanto que as características CorPolpa, *DMC.Grav*, AMD, *DMC.DO* e HCN apresentaram h^2 de magnitude mediana, ou seja, 0,42, 0,45, 0,48, 0,54 e 0,60, respectivamente. Por outro lado, as características produtivas foram as que apresentaram a menor herdabilidade ($h^2=0,32$ para NRP e IC, e h^2 entre 0,21 e 0,22 para PTR, *PROD.Dry* e PPA). Entretanto, a herdabilidade no sentido amplo por parcela foi bastante elevada para todas as características, especialmente para PTR, *PROD.Dry*, PPA, NRP e IC (h_m^2 entre 0,80 e 0,89).

3.2 Correlações entre as características

Na Figura 2 são apresentadas as correlações de Pearson, que indicaram fortes correlações positivas entre TCC x CorPolpa ($r=0,70$) e PTR x *PROD.Dry* ($r=0,93$) e correlações positivas de média magnitude entre *DMC.Grav* versus *DMC.OD* ($r=0,59$) e AMD ($r=0,58$), *DMC.OD* x AMD ($r=0,64$), PTR x PPA ($r=0,58$), e PPA x *PROD.Dry* ($r=0,53$). Além disso, correlações positivas de baixa magnitude, mas também significativas foram identificadas entre PTR versus IC ($r=0,47$) e NRP ($r=0,42$), IC versus *PROD.Dry* ($r=0,42$) e NRP ($r=0,32$), *PROD.Dry* x NRP ($r=0,35$) e CorPolpa x HCN ($r=0,29$). As correlações negativas observadas foram significativas, mas de baixa magnitude entre PPA x IC ($r=-0,36$), IC x HCN ($r=-0,32$) e NRP x HCN ($r=-0,29$). As correlações entre TCC versus *DMC.Grav* e *DMC.OD* foram negativas ($r=-0,04$ e $-0,05$, respectivamente) mas praticamente nulas e sem significância.

A *network* construída com base em todos os pares de correlações significativas entre atributos agronômicos e de qualidade de raízes de mandioca é apresentada na Figura 3. Como esperado, esta *network* de correlações mostrou tendências semelhantes ao correlograma da Figura 2. Além disso, essa *network* identificou uma clara separação entre as

características agronômicas e de qualidade de raízes de mandioca. Além disso, independentemente do sentido da correlação, houve uma relação muito forte entre as características agronômicas IC, PPA, PTR e *PROD.Dry*, enquanto a característica NRP possui uma menor associação positiva com os demais atributos agronômicos e negativa com a qualidade das raízes. Dentre as características agronômicas o IC e PTR foram as que apresentaram maior conexão com as demais características, enquanto PPA e *PROD.Dry* apresentaram menor número de conexões.

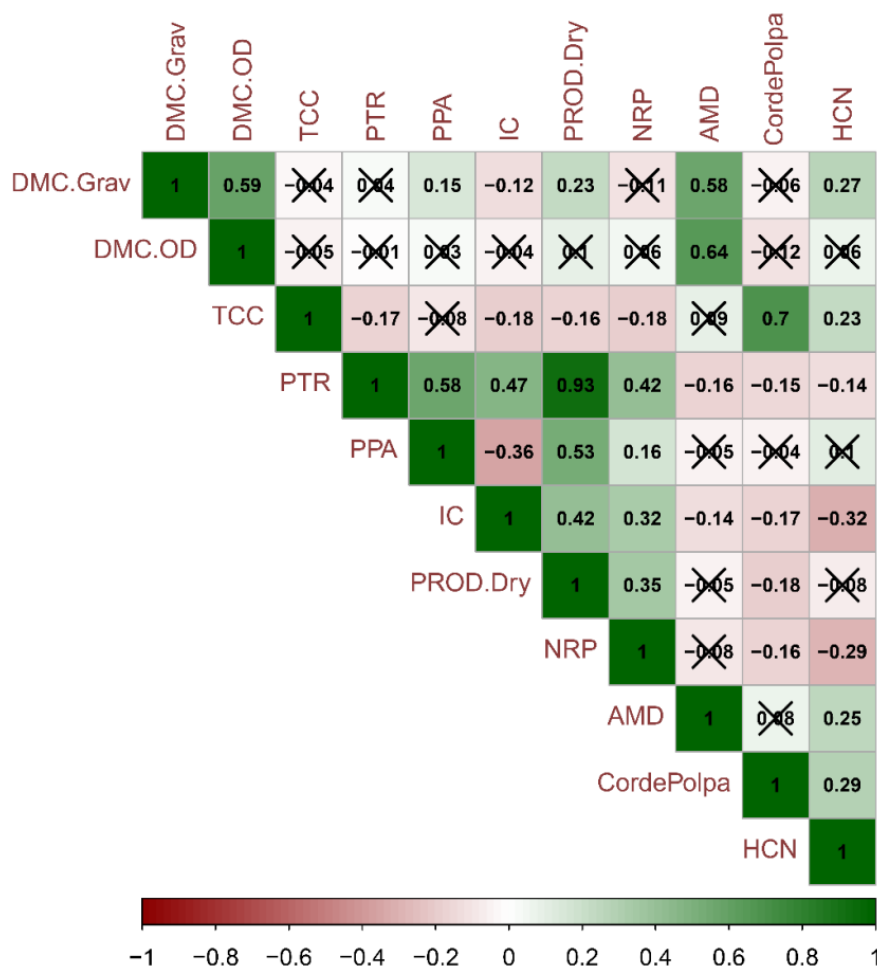


Figura 2. Correlograma de Pearson entre características agronômicas e de qualidade de raízes de mandioca avaliados em genótipos de mandioca de polpa amarela. TCC: teor de carotenoides totais; *DMC.Grav*: teor de matéria seca pelo método gravimétrico e *DMC.DO* pelo método da secagem; PTR: produtividade de raízes frescas; PPA: produtividade da parte aérea; IC: índice de colheita; *PROD.Dry*: produtividade de raízes secas; NRP: número de raízes; AMD: teor de amido; CorPolpa: cor da polpa das raízes; HCN: teor de compostos cianogênicos. As correlações marcadas com “x” não são significativas ($p > 0,05$).

Para a qualidade de raízes, houve a formação de dois subgrupos de características sendo *DMC.Grav*, *DMC.DO* e *AMD* em um subgrupo e *HCN*, *TCC* e *CorPolpa* em outro subgrupo. Apesar das baixas correlações (negativas e positivas), *HCN* e *DMC.Grav* tiveram maior número de conexões com as demais características. Por outro lado, *TCC* relacionou-se basicamente com o *CorPolpa* e alguns atributos produtivos em baixa correlação.

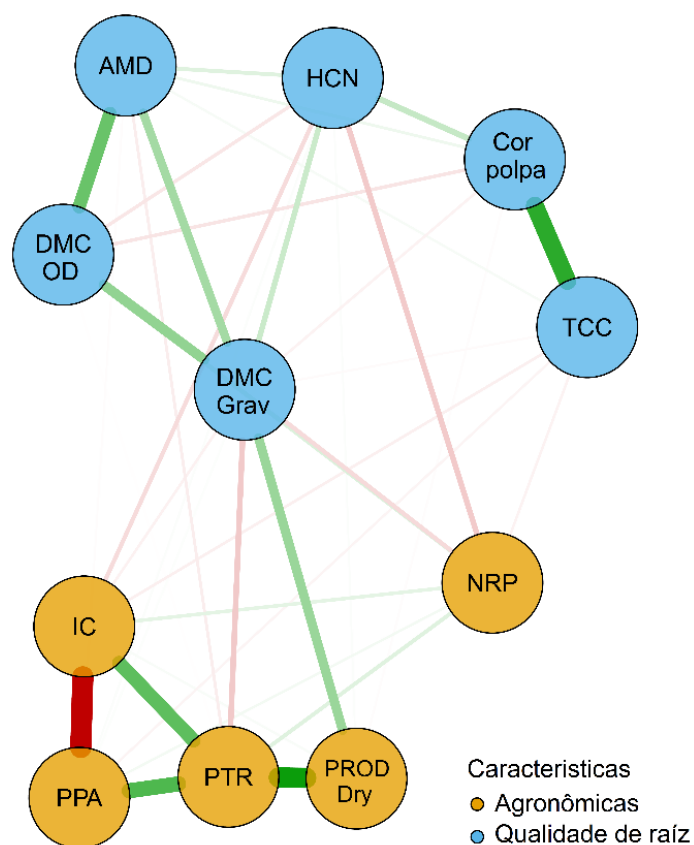


Figura 3. Rede regularizada das correlações parciais entre características agronômica e de qualidade de raízes de mandioca avaliados em genótipos de mandioca de polpa amarela. A espessura das retas representa as correlações, enquanto as cores verde e vermelha representam as correlações positivas e negativas, respectivamente. *TCC*: teor de carotenoides totais; *DMC.Grav*: teor de matéria seca pelo método gravimétrico; *DMC.DO*: teor de matéria seca pelo método da secagem em estufa; *PTR*: produtividade de raízes frescas; *PPA*: produtividade da parte aérea; *IC*: índice de colheita; *PROD.Dry*: produtividade de raízes secas; *NRP*: número de raízes; *AMD*: teor de amido; *CorPolpa*: cor da polpa das raízes; *HCN*: teor de compostos cianogênicos.

3.3. Análise do padrão de diversidade dos clones de polpa amarela

O critério de informação bayesiano (BIC) indicou a formação de seis grupos distintos de diversidade das características associadas à produtividade e qualidade das raízes no germoplasma de mandioca de polpa amarela (Figura 4). Houve uma distribuição similar dos acessos em cada agrupamento, ou seja, 24, 60, 38, 39, 60 e 44 genótipos nos clusters 1, 2, 3, 4, 5 e 6, respectivamente. No cluster1 foram alocadas 17 variedades locais e sete variedades melhoradas, incluindo duas biofortificadas lançadas no mercado (BRS Jari, BRS-396 e BRS-399). No cluster2 foram agrupadas 45 variedades locais e 15 variedades melhoradas, incluindo BRS Dourada e BRS Gema de Ovo, que também foram recomendadas como variedades com alto teor de carotenoides para uso como mandioca de mesa. Dos 38 genótipos do cluster3, 11 são variedades melhoradas que na maioria dos casos foram recomendadas para a região amazônica, enquanto que as demais 27 variedades locais foram coletadas em diferentes regiões do Brasil. No caso do cluster4, apenas seis dos 39 genótipos são classificados como variedades melhoradas (muitas delas sejam obsoletas), enquanto mais de 60% das 33 variedades locais deste grupo foram coletadas na região Nordeste do Brasil. O cluster5 possui apenas seis variedades melhoradas (com cultivo restrito no Brasil) e 54 variedades locais, sendo 20 delas oriundas da região Norte, 23 da região Nordeste do Brasil e o restante de origem brasileira, porém sem identificação precisa da região de coleta. Por último, sete dos 44 genótipos do cluster6 são variedades melhoradas para consumo de mesa, enquanto as demais variedades locais tem origem de coleta e tradição de uso na região Norte (13 genótipos) e Nordeste (18 genótipos). As demais variedades locais do cluster6 não tem origem definida.

As características do cluster1 foram o baixo teor de matéria seca nas raízes (30,2% e 31,5% pelos métodos gravimétrico e secagem em estufa, respectivamente), teores intermediários de TCC ($\sim 4,5 \mu\text{g.g}^{-1}$) associado à nota média de 2,35 para cor de polpa, indicando que a maioria dos genótipos possui raízes de cor creme e valores intermediários de HCN (6,6). Além disso, disso os genótipos deste cluster possuem baixo AMD ($\sim 21,5 \%$) e baixo potencial produtivo: PTR ($\sim 14,2 \text{ t.ha}^{-1}$), PPA ($\sim 16,0 \text{ t.ha}^{-1}$), *PROD.Dry* ($\sim 3,9 \text{ t.ha}^{-1}$) e NRP ($\sim 3,6$) (Figuras 5 e 6).

características estão numeradas de 1 a 11, na seguinte ordem: teor de compostos cianogênicos, cor da polpa das raízes, teor de amido, número de raízes por planta, produtividade de raízes secas, índice de colheita, produtividade da parte aérea, produtividade de raízes frescas, teor de carotenoides totais, teor de matéria seca pelo método da secagem em estufa e teor de matéria seca pelo método gravimétrico.

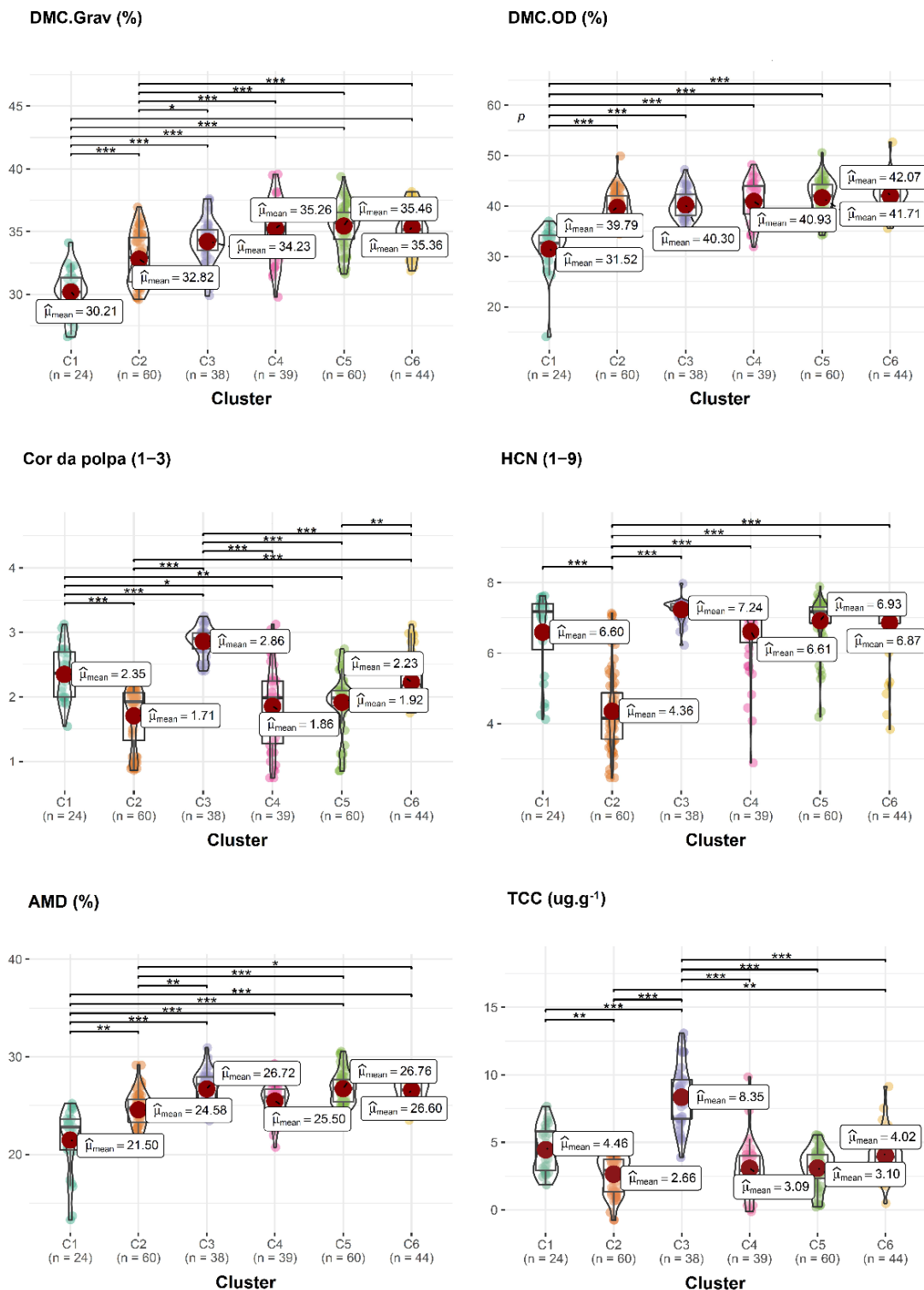


Figura 5. Box/Violino plot dos *best linear unbiased prediction* (BLUP) dos diferentes agrupamentos dos genótipos de mandioca de polpa amarela para as características relacionadas à qualidade de raízes: *DMC.Grav*: teor de matéria seca pelo método gravimétrico, *DMC.DO*: teor de matéria seca pelo método da secagem em estufa, *CorPolpa*: cor da polpa das raízes, *HCN*: teor de compostos cianogênicos, *AMD*: teor de amido, e *TCC*: teor de carotenoides totais. *, **, *** se refere a comparações significativas pelo método Holm a 0,05, 0,01 e 0,001, respectivamente.

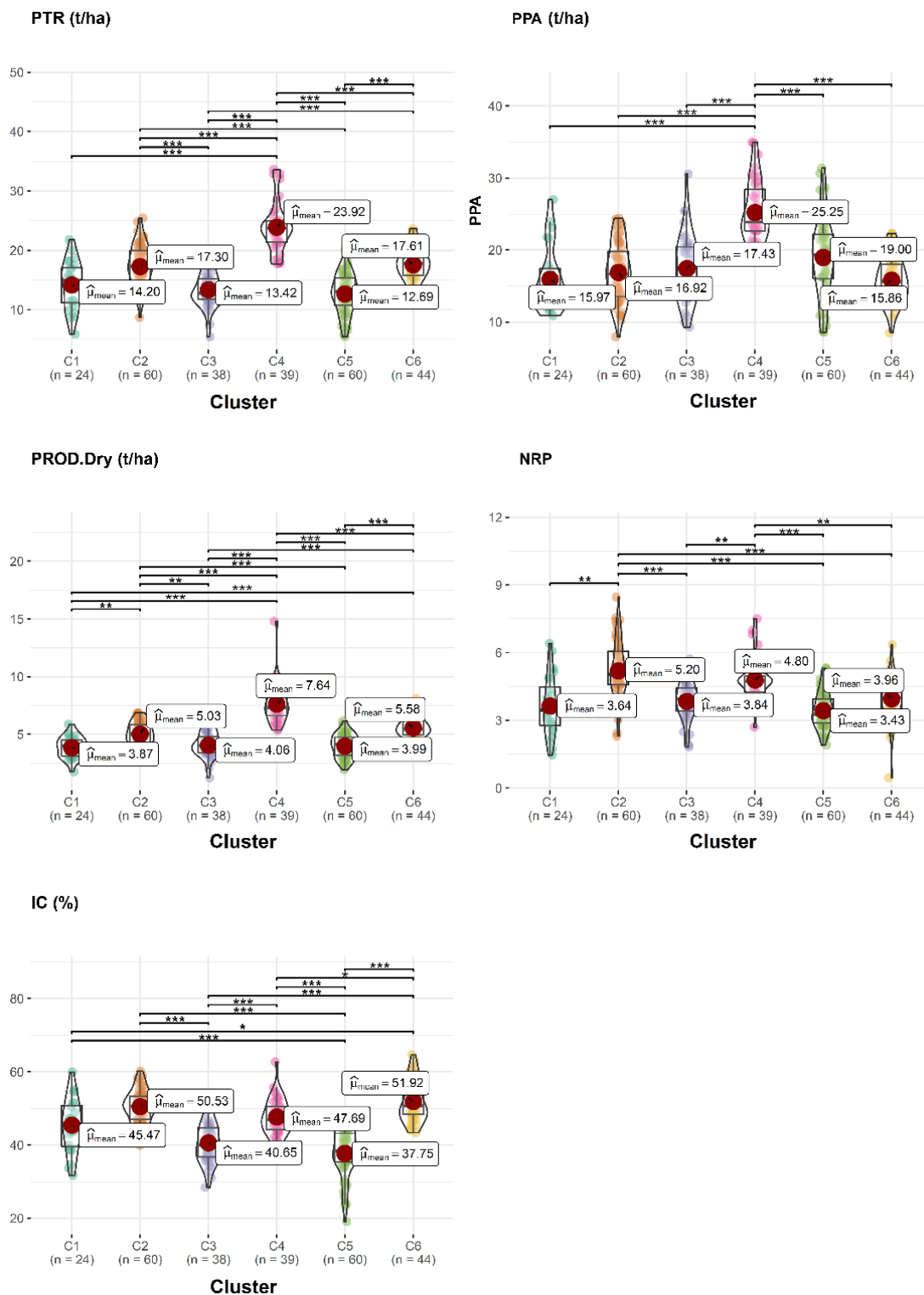


Figura 6. Box/Violino plot dos *best linear unbiased prediction* (BLUP) dos diferentes agrupamentos dos genótipos de mandioca de polpa amarela para as características agrônômicas: PTR: produtividade de raízes frescas, PPA: produtividade da parte aérea; PROD.Dry: produtividade de raízes secas, NRP: número de raízes e IC: índice de colheita. *, **, *** se refere a comparações significativas pelo método Holm a 0,05, 0,01 e 0,001, respectivamente.

Em relação às características associadas à qualidade de raízes de mandioca o cluster3 apresentou teores intermediários de matéria seca nas raízes (34,23% e 40,30% pelos métodos gravimétrico e secagem em estufa, respectivamente), e maior média para TCC ($\sim 8,3 \mu\text{g.g}^{-1}$) associado à cor da polpa com nota média igual a 2,86 (maioria dos clones de polpa amarela) e elevado HCN (7,24). Por outro lado, os genótipos do cluster3 possuem baixo potencial produtivo, sobretudo para PTR ($\sim 13,4 \text{ t.ha}^{-1}$), PPA ($\sim 17,4 \text{ t.ha}^{-1}$), IC ($\sim 40,6 \%$), *PROD.Dry* ($\sim 4,1 \text{ t.ha}^{-1}$) e NRP ($\sim 3,8$).

De forma semelhante ao cluster2, o cluster4 apresentou baixo TCC ($\sim 3,1 \mu\text{g.g}^{-1}$) e cor da polpa (score $\sim 1,9$) indicando que existem alguns clones com polpa branca embora a maioria dos genótipos polpa creme. Entretanto o cluster4 possui valores medianos de AMD ($\sim 25,5 \%$) e HCN (6,6), bem como altos teores de matéria seca pelos métodos gravimétrico e secagem em estufa ($\sim 35,3\%$ e $\sim 41,0\%$, respectivamente). Além disso, o cluster4 foi representado pelos genótipos com a maior PTR ($\sim 24,0 \text{ t.ha}^{-1}$), PPA ($\sim 25,3 \text{ t.ha}^{-1}$), *PROD.Dry* ($\sim 7,6 \text{ t.ha}^{-1}$) e NRP ($\sim 4,8$).

Os atributos do cluster5 foram os altos teores de matéria seca nas raízes (35,5% e 41,7% pelos métodos gravimétrico e secagem em estufa, respectivamente) e AMD ($\sim 26,8\%$) e HCN (6,9). Por outro lado, os teores de carotenoides e a cor das raízes também foram baixos ($\sim 3,1 \mu\text{g.g}^{-1}$ e 1,92, respectivamente). Em termos agronômicos os genótipos do cluster5 possuem baixo NRP ($\sim 3,4$) associado à baixa PTR ($\sim 12,7 \text{ t.ha}^{-1}$) e *PROD.Dry* ($\sim 3,9 \text{ t.ha}^{-1}$), embora tenham apresentado alto PPA ($\sim 19,0 \text{ t.ha}^{-1}$) e consequentemente o menor IC ($\sim 37,7\%$), quando comparados aos demais clusters.

Por fim, juntamente com o cluster5, o cluster6 apresentou os mais elevados teores de matéria seca nas raízes pelo método gravimétrico ($\sim 35,3\%$) e secagem em estufa ($\sim 42,1\%$), bem como elevado AMD ($\sim 26,6\%$) e HCN (6,9). Os genótipos do cluster6 possuem valores medianos de TCC ($\sim 4,0 \mu\text{g.g}^{-1}$) e cor da polpa (2,2), indicando a presença de um mix de clones de cor creme e amarela neste grupo. Uma outra característica marcante deste grupo é a presença de clones com alto IC ($\sim 52\%$), como resultado do maior peso da PTR ($\sim 17,6 \text{ t.ha}^{-1}$) em comparação com a PPA ($\sim 15,9 \text{ t.ha}^{-1}$). Os valores de

PROD.Dry ($\sim 5,6 \text{ t.ha}^{-1}$) e NRP ($\sim 4,0$) foram considerados de magnitude mediana em comparação com os demais grupos.

3.4. Seleção de genótipos

A seleção dos 30 melhores genótipos para recombinação no programa de melhoramento de mandioca tem potencial para elevar o TCC em 37% quando comparado com a média geral do germoplasma, chegando a $\sim 5,6 \mu\text{g.g}^{-1}$ (Tabela 3), associado com uma forte redução no HCN (-19,4%). Além disso, existe potencial de ganhos expressivos para diversos atributos agrônômicos como PTR (37,7%), PPA (21%), PROD.Dry (40%) e NRP (17,4%). Por outro lado, os ganhos para teor de matéria seca nas raízes são potencialmente pequenos (1,9% e 2,6% considerando os métodos gravimétrico e da secagem, respectivamente).

Tabela 3. Lista de genótipos selecionados para recombinação visando a geração de progênies biofortificadas de mandioca de mesa com base no índice de seleção, bem como a média geral do germoplasma avaliado e clones selecionados e ganhos potenciais.

Clone	Cluster	DMC.Grav	DMC.OD	TCC	PTR	PPA	IC	PROD.Dry	NRP	AMD	CordePolpa	HCN
Folha-Fina	4	35,91	37,84	3,32	32,86	31,47	47,96	9,79	4,91	25,03	2,07	2,89
BGM-1390	4	38,18	41,59	4,56	26,90	35,00	46,95	9,79	6,99	25,85	3,04	7,33
BGM-1962	4	35,30	39,15	9,85	24,31	23,77	47,32	7,78	4,09	25,54	2,68	7,40
BGM-0597	3	32,80	42,16	13,08	17,94	21,97	42,99	4,82	4,71	25,99	2,86	7,26
BGM-2127	4	33,40	36,44	3,26	32,90	18,61	62,69	14,81	4,75	21,97	1,14	5,49
CPAFRO-04	6	38,01	42,73	7,53	23,67	18,00	57,82	7,09	3,92	26,35	2,99	7,46
BGM-0019	3	36,05	44,10	11,67	17,15	20,47	42,90	5,30	4,26	28,87	2,74	7,23
BGM-1692	4	32,25	39,40	4,51	24,91	23,61	49,85	6,97	4,99	24,32	2,64	4,08
BGM-1700	6	34,81	39,98	9,14	20,06	20,00	49,04	6,46	4,21	25,89	1,91	6,15
BGM-1709	3	33,43	41,43	9,15	20,20	20,90	46,59	5,99	5,72	26,25	2,91	7,38
BGM-0290	4	38,06	43,36	2,57	23,95	23,03	50,40	7,74	3,42	24,70	2,98	4,45
BGM-1835	2	35,09	42,13	2,22	22,02	17,57	56,21	6,86	7,40	27,42	1,94	3,50
BGM-0368	2	33,35	38,18	3,56	21,17	16,55	53,80	5,92	6,11	26,04	2,09	2,46
BGM-2353	6	38,19	43,11	5,88	16,79	16,44	48,88	7,31	2,45	27,99	2,14	3,84
BGM-1745	3	37,51	42,74	8,96	16,08	14,21	50,83	5,51	3,62	29,10	2,94	7,40
BGM-0918	4	36,15	43,11	2,09	32,06	33,27	41,18	10,17	4,23	25,90	1,37	7,08
BGM-0444	2	32,44	37,42	4,61	23,25	22,20	51,45	6,46	3,50	23,93	2,07	3,95

BGM-0901 4 36,40 46,12 2,60 28,41 29,28 46,15 8,73 4,66 26,18 1,85 7,46

Tabela 3. Cont....

Clone	Cluster	DMC.Grav	DMC.OD	TCC	PTR	PPA	IC	PROD.Dry	NRP	AMD	CordePolpa	HCN
BGM-1780	3	37,62	41,61	8,98	15,85	14,44	49,04	5,34	4,20	27,38	3,18	7,40
BGM-0212	2	34,03	44,02	5,50	16,01	16,66	48,87	4,63	4,83	25,92	2,14	2,81
BGM-1814	4	34,47	41,05	3,66	22,15	24,35	48,90	6,51	4,62	28,93	2,00	4,83
BGM-0579	4	36,64	45,66	7,34	21,31	24,46	41,82	6,97	4,85	26,76	1,25	7,32
BRS-Dourada	2	31,48	41,94	2,65	24,78	24,28	50,30	6,40	5,79	22,18	2,29	4,18
BGM-0120	4	31,40	34,11	2,07	33,62	34,84	50,70	8,90	5,19	22,67	1,80	7,32
BGM-0659	4	29,77	39,44	5,28	22,97	30,26	42,20	6,11	6,82	23,44	2,25	5,63
BGM-0760	4	32,22	38,93	4,09	25,79	18,32	54,34	7,18	4,72	24,70	2,25	5,63
BRS-Gema-de-Ovo	2	35,75	42,73	3,67	17,90	20,99	46,57	5,61	5,74	25,77	2,31	3,08
BGM-0936	3	33,93	39,58	9,68	16,19	25,49	35,37	5,38	5,00	28,50	3,00	7,41
BGM-1497	2	34,87	35,46	3,56	20,74	16,32	57,14	6,55	3,66	23,25	2,09	2,56
BGM-2345	2	36,09	44,85	3,15	19,21	17,59	52,97	5,75	6,78	23,30	1,89	3,57
Média sel		34,85	41,01	5,61	22,71	22,48	49,04	7,09	4,87	25,67	2,29	5,48
Média geral		34,22	39,99	4,09	16,49	18,57	45,40	5,06	4,15	25,36	2,11	6,80
Ganho potencial (%)		1,85	2,56	37,08	37,69	21,05	8,02	40,22	17,38	1,22	8,70	-19,35

Teor de matéria seca pelo método gravimétrico (*DMC.Grav*) e da secagem em estufa (*DMC.DO*), teor de carotenoides totais (TCC), produtividade de raízes frescas (PTR) e secas (*PROD.Dry*), produtividade da parte aérea (PPA), índice de colheita (IC), número de raízes (NRP), teor de amido (AMD), teor de compostos cianogênicos (HCN) e cor da polpa das raízes (CorPolpa).

A maioria dos genótipos selecionados para recombinação pertencem ao cluster4 (43,3%) e cluster2 (26,7%), cuja média das características produtivas e do baixo HCN foram preponderantes para a escolha dos indivíduos. Entretanto, também foram selecionados genótipos dos cluster3 (20%) e cluster6 (10%) que possuem principalmente alto TCC e DMC.

4. DISCUSSÃO

4.1. Variabilidade do germoplasma de mandioca de polpa amarela

Combinar qualidade nutricional com características que definem a aceitação do consumidor, como teor de matéria seca e potencial cianogênico, tem sido uma das principais preocupações no desenvolvimento de variedades biofortificadas de mandioca. O aumento do TCC em raízes de mandioca e os fatores envolvidos neste processo tem sido foco constante dos programas de

melhoramento de mandioca (ESUMA et al., 2012; CEBALLOS et al., 2013; RABBI et al., 2017; CEBALLOS et al., 2017; BEYENE et al., 2017). Por meio de seleção recorrente, Ceballos et al. (2013) conseguiram elevar o TCC de 10,3 para 24,3 $\mu\text{g g}^{-1}$ em um período de nove anos. Esse aumento é possível devido à predominância de efeitos genéticos aditivos envolvidos na herança do TCC em mandioca (ESUMA et al., 2016) e que, portanto, o melhoramento populacional pode ser bastante efetivo na recombinação dos alelos favoráveis e consequente aumento no TCC. Para um programa de melhoramento atingir tais objetivos, a exploração da variabilidade genética presente no banco de germoplasma é uma abordagem inicial para a identificação de genótipos com elevados valores de carotenoides aliados a outras características de interesse.

Nesse estudo foi realizado um *screening* do germoplasma de mandioca de polpa amarela do Brasil, para avaliar a variabilidade quanto ao teor de carotenoides e de características relacionadas a qualidade de raízes e ao potencial agrônômico. Os genótipos foram escolhidos por apresentarem algum nível de pigmentação amarela na raiz, tendo em vista a alta correlação entre cor da polpa \times TCC (CHÁVEZ et al. 2005; SÁNCHEZ et al., 2006; ESUMA et al., 2016). Os valores de TCC deste estudo baseado em 265 genótipos apresentaram maiores amplitudes (0,075 $\mu\text{g g}^{-1}$ a 13,08 $\mu\text{g g}^{-1}$) em relação ao relato de Chávez et al. (2005) (1,02 a 10,40 $\mu\text{g g}^{-1}$) em 1789 genótipos de diferentes origens e também com diferentes níveis de melhoramento. Por outro lado, a amplitude de variação da característica *DMC.Grav* (10,72 a 57,23%) identificada em 2022 genótipos foi mais elevada em comparação ao presente estudo (26,61 a 39,57%), embora o *DMC.OD* tenha apresentado amplitude muito semelhante (14,07 a 52,67%). Essa amplitude para *DMC.OD* foi superior aos achados de Rabbi et al. (2017), que identificaram variação de 8,4 a 45,4% baseada em 3232 genótipos desenvolvidos no Instituto Internacional de Agricultura Tropical (IITA).

A estimativa dos valores de PTR também possui grande importância para o ranqueamento, seleção e adoção de variedade de mandioca. No presente estudo a média geral de 16,49 t.ha^{-1} de raízes frescas foram semelhantes ao relato de Parkes et al. (2020), ao analisarem genótipos parentais elite de mandioca amarela do IITA (média de 16,32 t.ha^{-1}).

Os genótipos avaliados nesse estudo, apresentam diferentes origens de coleta e níveis de melhoramento (Tabela suplementar 1), sendo um provável motivo para a presença da ampla variabilidade genética identificada (AVIJALA et al., 2015). As avaliações de características de interesse agrônômico e de qualidade de raiz, permitem identificar parentais com potencial para o desenvolvimento de progênies segregantes com máxima variabilidade genética e introgressão de genes desejáveis disponíveis no germoplasma (OLIVEIRA et al., 2014).

4.2. Parâmetros genéticos e correlações entre atributos produtivos e de qualidade de raízes

As estimativas de herdabilidade em sentido amplo são classificadas como baixas ($<0,30$), moderadas ($0,30$ a $0,60$) e altas ($>0,60$) de acordo com Mehari; Amsalu; Tewodros, (2015). Altas estimativas de herdabilidade no sentido amplo foram encontradas para TCC ($h^2=0,72$). Valores similares para h^2 da característica TCC foram relatados em outros estudos, a exemplo de Esuma et al. (2012) ($h^2=0,73$) e Parkes et al. (2020) ($h^2=0,60$). A ausência do efeito da interação genótipo \times ambiente e alta herdabilidade para teor de carotenoides sugere a possibilidade de resposta à seleção via recombinação dos alelos favoráveis. Além disso, valores elevados de herdabilidade permitem uma seleção de parentais mais acurada (PARKES et al., 2020; CEBALLOS et al., 2017). Por outro lado, o teor de matéria seca apresentou h^2 de magnitude mediana (DMC.Grav $0,45$ e DMC.DO $0,54$), sendo similares aos valores reportados por Parkes et al. (2020) e Rabbi et al. (2017), que também verificaram valores de h^2 moderados (variação de $0,40$ a $0,51$) para esta característica.

As características produtivas PTR ($h^2 =0,21$), *PROD.Dry* ($h^2 =0,21$) e PPA ($h^2 =0,22$) foram de baixa magnitude. Estes valores tendem a dificultar o processo de seleção para o melhoramento genético, devido à maior influência do ambiente (CEBALLOS et al., 2004). Uma alternativa para o melhoramento de características com este tipo de herdabilidade seria submeter a população a métodos de melhoramento populacional como seleção fenotípica recorrente visando aumentar a frequência de alelos favoráveis ao longo dos ciclos de

seleção. O objetivo deste tipo de seleção é obter melhoramento populacional ao longo de diferentes ciclos de seleção, pelo acúmulo de alelos favoráveis (BOS; CALIGARI, 2007).

As correlações entre TCC e HCN foram positivas, mas de baixa magnitude. Portanto, é possível identificar genótipos com alto conteúdo de TCC nas raízes e que ao mesmo tempo possam ser utilizadas como variedades de mesa (HCN baixo ou intermediário). Variedades biofortificadas de mandioca devem ser utilizadas para consumo “*in natura*” porque o cozimento das raízes retém entre 95-100% do teor de carotenoides, enquanto que o processamento das raízes retém entre 2-5% na forma de fufu (massa cozida de mandioca) e entre 26-29% na forma de chikwangué (massa de mandioca embebida, enxaguada, amassada e cozida no vapor dentro de folhas, normalmente de bananeira) (TALEON et al., 2019). Portanto, é necessário que as raízes possuam ≤ 100 mg de compostos cianogênicos kg^{-1} (PEPRAH et al., 2020), limite aceitável de cianeto para consumo humano (NHASSICO et al., 2008; FALADE; AKINGBALA, 2010; CLIFF et al., 2011).

A correlação não significativa entre TCC \times DMC (*DMC.Grav* e *DMC.OD*) no germoplasma de mandioca biofortificada no Brasil permite que o melhoramento da espécie selecione indivíduos com alto TCC e DMC de forma simultânea, para atender à demanda dos usuários finais das variedades. O DMC em variedades biofortificadas auxilia na retenção de carotenoides após o cozimento, evitando que estes pigmentos fluam para água durante o processo (CEBALLOS et al., 2012) e portanto, as variedades que combinem essas duas características tem maior probabilidade de recomendação de cultivo (NJUKWE et al. 2013; PEPRAH et al., 2020).

Estudos prévios demonstraram que as características TCC e DMC são herdados de forma independente (SÁNCHEZ et al., 2006). Em um estudo mais recente conduzido por Sánchez et al. (2014), utilizando amostras de germoplasma da América do Sul, os autores reportaram ausência de correlação entre TCC versus DMC. Entretanto, a análise de um painel de germoplasma oriundos da África (clones melhorados e variedades locais), demonstrou a existência de correlações negativas entre essas características em raízes de mandioca, com valores variando entre -0,22 a -0,59 (AKINWALE

et al., 2010; ESUMA et al., 2012; NJOKU et al., 2015; RABBI et al., 2017). A existência desta correlação negativa no germoplasma africano de mandioca tem sido atribuída à presença de dois loci no cromossomo 1 nas posições 24.1 e 30.5 Mbp, associados à cor das raízes de mandioca (altamente correlacionada com teor de carotenoides) e um único locus para o teor de matéria seca na região próxima a 24,1 Mbp no cromossomo 1, ocasionando uma ligação física entre estas duas características no germoplasma de mandioca africano (RABBI et al., 2017).

Embora a presença de correlação negativa entre TCC × DMC seja indesejável para o desenvolvimento de variedades melhoradas, é possível proceder sucessivas recombinações em esquema de seleção recorrente para quebrar essa ligação física entre essas características. O uso deste tipo de abordagem tem sido apresentado como caso de sucesso na obtenção de genótipos com elevados valores de TCC e com DMC em níveis aceitáveis, ao longo de nove ciclos de seleção, elevando o TCC em mais de 100% em comparação com a população original (CEBALLOS et al., 2013; SÁNCHEZ et al., 2014).

Além do germoplasma africano, relatos da existência de correlação negativa entre TCC × DMC também foram relatados em variedades transgênicas, pois o aumento no TCC resultou em efeitos pleiotropicos indesejados para redução do DMC (FAILLA et al., 2012). Beyene et al. (2017) reportaram uma forte correlação negativa entre TCC × DMC em raízes de mandioca transgênica cultivadas em casa-de-vegetação ($r = -0,85$, $P \leq 0,001$) e em campo ($r = -0,92$, $P \leq 0,001$) em comparação com controles não transgênicos. As principais razões apontadas pelos autores para essa correlação negativa é a mudança no fluxo metabólico observado nas raízes modificadas, em que o desvio do piruvato para a síntese de carotenoides possa ter suprimido a atividade da ADP-glicose pirofosforilase (enzima importante para a biossíntese do amido).

A existência de germoplasma de mandioca nativo não transgênico no Brasil com correlação nula entre TCC × DMC e ao mesmo tempo com elevada herdabilidade abre importantes possibilidades de recombinação de alelos úteis para estas duas características utilizando abordagens simples de

melhoramento, à exemplo da seleção fenotípica. Entretanto, o uso de abordagens mais sofisticadas como a seleção genômica tem demonstrado alto potencial para rápida incorporação de alelos úteis e aumento destas características de forma simultânea (ESUMA et al., 2021). Esses autores avaliaram a capacidade de predição genômica dentro da população de base para carotenoides e identificaram a possibilidade de predizer TCC e DMC no germoplasma da África Ocidental usando esta população, com acurácia média de 0,40 e 0,34, respectivamente, utilizando o modelo de predição G-BLUP.

4.3. Seleção de genótipos de mandioca de polpa amarela

A seleção dos 30 genótipos para recombinação no programa de melhoramento de mandioca foi baseada no alto TCC e baixo teor de HCN, associados a alta performance agrônômica. Dentre esses genótipos estão algumas variedades melhoradas do programa de biofortificação da Embrapa (BRS Dourada e BRS Gema de Ovo) por possuírem valores medianos de TCC ($3,16 \mu\text{g.g}^{-1}$) associados com alto potencial produtivo ($21,34 \text{ t.ha}^{-1}$ PTR, 42,33 % *DMC.OD* e $6,0 \text{ t.ha}^{-1}$ *PROD.Dry*). Além disso, os genótipos BGM-0659, BGM-0212 e BGM-2353 apresentaram valores de TCC superior às variedades melhoradas ($5,55 \mu\text{g.g}^{-1}$), com bom potencial produtivo ($18,59 \text{ t.ha}^{-1}$ PTR, 42,19 % *DMC.OD*, $6,01 \text{ t.ha}^{-1}$ *PROD.Dry*), associados com valores de HCN considerados que podem ser considerados baixos (2,81 e 3,84 para os genótipos BGM-0212 e BGM-2353, respectivamente) e intermediários (5,63 para o genótipo BGM-0659). Todos estes parentais promissores irão compor blocos de cruzamentos para geração de progênes com alta variabilidade para essas características, tendo-se o cuidado para garantir que as características agrônômicas da cultura, como produtividade, não fossem afetadas negativamente. Isso pode ser comprovado com base no ganho genético potencial para atributos produtivos como *PROD.DRY* (40,22%), PTR (37,69%) e PPA (21,05%). Portanto, este painel de parentais tem elevado potencial para gerar progênes superiores e transgressivas.

A importância primária em desenvolver essas variedades com elevado TCC é por enriquecer nutricionalmente a dieta, a partir do fornecimento do β -caroteno como pró-vitamina A, por meio de suas raízes. Neste sentido, espera-

se que nos próximos ciclos de reprodução sejam desenvolvidas variedades de mandioca de mesa com níveis elevados do conteúdo de carotenoides, ajustados com características agronômicas, assim como características organolépticas e que apresentem estabilidade de cultivo em ambientes diversos.

5. CONCLUSÃO

A ampla variabilidade genética foi identificada no germoplasma de mandioca de polpa amarela para atributos associados à qualidade das raízes e de performance agronômica e o agrupamento dos genótipos em seis grupos bastante distintos quanto aos atributos avaliados, constitui-se em um importante ponto de partida para posteriores estudos de conservação e uso destes recursos genéticos para fins de melhoramento.

As magnitudes e direções das correlações identificadas entre as características de maior impacto agronômico no desenvolvimento de variedades de mandioca de mesa certamente contribuirá para a elaboração de melhores estratégias de seleção e composição de pesos econômicos em índices de seleção.

A maior herdabilidade para TCC ($h^2=0,72$) em raízes de mandioca, indica que a seleção de plantas com esta característica pode ser iniciada em fases mais precoces do programa de melhoramento na qual existem clones não replicados e cultivados em apenas um ambiente, a exemplo da fase de avaliação de seedlings e ensaio de avaliação clonal. Por outro lado, para as características de herdabilidade baixa ou mediana a exemplo de DMC, a seleção deve ser iniciada nos ensaios que permitem a avaliação com repetição de clones dentro do ensaio e em diferentes ambientes, à exemplo dos ensaios preliminares, avançados e regionais de produção.

Os 30 genótipos selecionados como parentais para recombinação no programa de mandioca biofortificadas de mesa permitirão a obtenção de ganhos genéticos importantes e simultâneos, sobretudo para TCC (37,08%), DMC (2,56%), PTR (37,69%), PROD.Dry (40,22%) e HCN (-19,35%).

6. REFERÊNCIAS BIBLIOGRÁFICAS

AKINWALE, M. G. B.; ALADESANWA, R. D.; AKINYELE, B. O.; DIXON, A. G. O.; ODIYI, A. C. Inheritance of β -carotene in cassava (*Manihot esculenta* Crantz). **International Journal of Genetics and Molecular Biology**, v. 2, n. 10, p. 198-201, 2010. <https://doi.org/10.5897/IJGMB.9000034>.

ALAMU, E. O.; MAZIYA-DIXON, B.; FELDE, T. Z.; KULAKOW, P.; PARKES, E. Application of near infrared reflectance spectroscopy in screening of fresh cassava (*Manihot esculenta* Crantz) storage roots for provitamin A carotenoids. Proceedings of the 18th International Conference on Near Infrared Spectroscopy. Chichester, **IM Publications**, p. 91–97, 2019. <https://doi.org/10.1255/nir2017.091>.

AVIJALA, M. F.; BHERING, L. L.; PEIXOTO, L. DE A.; CRUZ, C. D.; CARNEIRO, P. C. S.; CUAMBE, C. E.; ZACARIAS, A. Evaluation of cassava (*Manihot esculenta* Crantz) genotypes reveals great genetic variability and potential selection gain. **Australian Journal of Crop Science**, v. 9, n. 10, p. 940-947, 2015. <https://search.informit.org/doi/10.3316/informit.659853541582368>

BALAGOPALAN, C. Cassava utilization in food, feed and industry. **Cassava: Biology, production and utilization**, 2002, 332p.

BATES, D.; MÄCHLER, M.; BOLKER, B.; WALKER, S. “Fitting Linear Mixed-Effects Models Using lme4.” **Journal of Statistical Software**, v. 67, n.1, p. 1-48, 2015. doi: 10.18637/jss.v067.i01.

BAYATA, A; GURE, A.; KEDIR, K. Evaluation of Nutritional Compositions of Cassava Roots. **Food Science and Nutrition Completed Research**, p. 177-187, 2018. doi:2000-02-06T21:52:52Z.

BELALCAZAR, J.; DUFOUR, D.; ANDERSSON, M. S.; PIZARRO, M.; LUNA, J.; LONDOÑO, L.; MORANTE, N.; JARAMILLO, A. M.; PINO, L.; LÓPEZ-LAVALLE, B. L. A.; DAVRIEUX, F.; TALSMA, E. F.; CEBALLOS, H. High-throughput phenotyping and improvements in breeding cassava for increased carotenoids in the roots. **Crop Science**, v. 56, n. 6, p. 2916-2925, 2016. <https://doi.org/10.2135/cropsci2015.11.0701>.

BEYENE, G.; SOLOMON, F. R.; CHAUHAN, R. D.; GAITÁN-SOLIS, E.; NARAYANAN, N.; GEHAN, J.; SIRITUNGA, D.; STEVENS, R. L.; JIFON, J.; ECK, J. V.; LINSLER, E.; GEHAN, M.; ILYAS, M.; FREGENE, M.; SAYRE, R. T.; ANDERSON, P.; TAYLOR, N. J.; CAHOON, E. B. Provitamin A biofortification of cassava enhances shelf life but reduces dry matter content of storage roots due to altered carbon partitioning into starch. **Plant biotechnology journal**, v. 16, n. 6, p. 1186-1200, 2017. <https://doi.org/10.1111/pbi.12862>.

BOS, I.; CALIGARI, P. **Selection methods in plant breeding**. Springer Science & Business Media, v.1, 2007, 461p.

BOUIS, H. E.; SALTZMAN, A. Improving nutrition through biofortification: a review of evidence from HarvestPlus, 2003 through 2016. **Global food security**, v. 12, p. 49-58, 2017. <https://doi.org/10.1016/j.gfs.2017.01.009>

BURNS, A.; GLEADOW, R.; CLIFF, J.; ZACARIAS, A.; CAVAGNARO, T. Cassava: the drought, war and famine crop in a changing world. **Sustainability**, v. 2, n. 11, p. 3572-3607, 2010. <https://doi.org/10.3390/su2113572>

CARMO, C. D.; SILVA, M. S.; OLIVEIRA, G. A. F.; OLIVEIRA, E. J. Molecular assisted selection for resistance to cassava mosaic disease in *Manihot esculenta* Crantz. **Scientia Agricola**, v.72, p.520-527, 2015. <https://doi.org/10.1590/0103-9016-2014-0348>.

CEBALLOS, H. IGLESIAS, C. A.; PÉREZ, J. C.; DIXON, A. G. Cassava breeding: opportunities and challenges. **Plant molecular biology**, v. 56, n. 4, p. 503-516, 2004. <https://doi.org/10.1007/s11103-004-5010-5>.

CEBALLOS, H.; DAVRIEUX, F.; TALSMA, E. F.; BELALCAZAR, J.; CHAVARRIAGA, P.; ANDERSSON, M. S. Carotenoids in cassava roots. In: **Carotenoids**. InTech, p. 189-221, 2017. <http://dx.doi.org/10.5772/intechopen.68279>.

CEBALLOS, H.; LUNA, J.; ESCOBAR, A. F.; ORTIZ, D.; PÉREZ, J. C.; SÁNCHEZ, T.; PACHÓN, H.; DUFOUR, D. Spatial distribution of dry matter in yellow fleshed cassava roots and its influence on carotenoid retention upon boiling. **Food research international**, v. 45, n. 1, p. 52-59, 2012. <https://doi.org/10.1016/j.foodres.2011.10.001>.

CEBALLOS, H.; MORANTE, N.; SÁNCHEZ, T.; ORTIZ, D.; ARAGÓN, I.; CHÁVEZ, A. L.; PIZARRO, M.; CALLE, F.; DUFOUR, D. Rapid cycling recurrent selection for increased carotenoids content in cassava roots. **Crop Science**, v. 53, n. 6, p. 2342-2351, 2013. <https://doi.org/10.2135/cropsci2013.02.0123>

CHÁVEZ, A. L.; SÁNCHEZ, T.; JARAMILLO, G.; BEDOYA, J. M. I.; ECHEVERRY, J.; BOLAÑOS, E.A.; IGLESIAS, C. A. Variation of quality traits in cassava roots evaluated in landraces and improved clones. **Euphytica**, v. 143, p. 125–133, 2005. <https://doi.org/10.1007/s10681-005-3057-2>

CLIFF, J.; MUQUINGUE, H.; NHASSICO, D.; NZWALO, H.; BRADBURY, J. H. Konzo and continuing cyanide intoxication from cassava in Mozambique. **Food and Chemical Toxicology**, v. 49, n. 3, p. 631-635, 2011. doi: <https://doi.org/10.1016/j.fct.2010.06.056>.

CUEVAS, H. E.; SONG, H.; STAUB, J. E.; SIMON, P. W. Inheritance of β -carotene-associated flesh colour in cucumber (*Cucumis Sativus* L.) fruit.

Euphytica, v. 171, n. 3, p. 301-311, 2010. doi: <https://doi.org/10.1007/s10681-009-0017-2>.

EPSKAMP, S. Ggraph: network visualizations of relationships in psychometric data. **Journal of Statistical Software**, v. 48, n. 4, p.1-18, 2012.

ESUMA, W.; KAWUKI, R. S.; HERSELMAN, L.; LABUSCHAGNE, M. T. Diallel analysis of provitamin A carotenoid and dry matter content in cassava (*Manihot esculenta* Crantz). **Breeding science**, v.66, n.4, p. 15159, 2016. doi: <https://doi.org/10.1270/jsbbs.15159>.

ESUMA, W.; OZIMATI, A.; KULAKOW, P.; GORE, M. A.; WOLFE, M. D.; NUWAMANYA, E.; EGESI, C.; KAWUKI, R. S. Effectiveness of genomic selection for improving provitamin A carotenoid content and associated traits in cassava. **G3 Genes Genomes Genetics**, p. jkab160, 2021. <https://doi.org/10.1093/g3journal/jkab160>

ESUMA, W.; RUBAIHAYO, P.; PARIYO, A.; KAWUKI, R.; WANJALA, B.; NZUKI, I.; HARVEY, J. J. W.; BAGUMA, Y. Genetic diversity of provitamin A cassava in Uganda. **Journal of Plant Studies**, v. 1, n. 1, p. 60, 2012. doi:10.5539/jps.v1n1p60

FAILLA, M. L.; CHITCHUMROONCHOKCHAI, C.; SIRITUNGA, D.; DE MOURA, F. F.; FREGENE, M.; SAYRE, R. T. Retention during processing and bioaccessibility of b-carotene in high b-carotene transgenic cassava root. **Journal of Agricultural and Food Chemistry**, v. 60, p. 3861–3866, 2012. doi: <https://doi.org/10.1021/jf204958w>.

FALADE, K. O.; AKINGBALA, J. O. Utilization of cassava for food. **Food Reviews International**, v. 27, n. 1, p. 51-83, 2010. <https://doi.org/10.1080/87559129.2010.518296>.

FAOSTAT-FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS (FAO). Participação dos continentes na produção de mandioca em 2019. FAOSTAT Database Gateway – FAO. Disponível em: <http://www.fao.org/faostat/en/#data/QC>. Acesso em: 18 Maio.2021.

FOUNDATION AGRICULTURAL ORGANIZATION (FAO). Food Outlook: **BIANNUAL REPORT ON GLOBAL FOOD MARKETS**, 2018. Disponível em: <<http://www.fao.org/3/CA2320EN/ca2320en.pdf>>. Acesso em: 10 de Março de 2021.

FUHRMANN, E.; VIEIRA, E. A.; FIALHO, J. DE F.; FALEIRO, F. G.; CARVALHO, L. J. C. B. Agronomic performance and biochemical attributes of yellow-pulped elite sweet cassava clones. **Científica**, v. 47, n. 1, p. 77-82, 2019. <http://dx.doi.org/10.15361/1984-5529.2019v47n1p77-82>.

FUKUDA, W. M. G.; CARVALHO, H. W. L.; OLIVEIRA, L. A.; OLIVEIRA, I. R. D.; PINHO, J. L. N. DE; SANTOS, V. D. S.; NUTTI, M. R.; KIMURA, M. **BRS Jari**: Nova variedade de mandioca para mesa com alto teor de betacaroteno nas raízes. Cruz das Almas, BA: Embrapa Mandioca e Fruticultura Tropical, Aracaju, Sergipe: Embrapa Tabuleiros Costeiros, 2009.

FUKUDA, W. M. G.; FUKUDA, C.; DIAS, M. C.; XAVIER, J. J. B. N.; FIALHO, J. D. F. Cultivares de mandioca recomendadas para o Brasil-2002. **Embrapa Amazônia Ocidental-Circular Técnica (INFOTECA-E)**, 2002.

FUKUDA, W. M. G.; GUEVARA, C. L.; KAWUKI, R.; FERGUSON, M. E. (2010) **Selected morphological and agronomic descriptors for the characterization of cassava**. International Institute of Tropical Agriculture (IITA), Ibadan, Nigeria. 19 pp.

FUKUDA, W. M. G.; PEREIRA, M. E. C. **BRS Gema de Ovo**: mandioca de mesa biofortificada. Cruz das Almas, BA: Embrapa Mandioca e Fruticultura Tropical, 2005.

FUKUDA, W. M. G.; PEREIRA, M. E. C.; OLIVEIRA, L. A.; GODOY, R. C. B. **BRS Dourada**: mandioca de mesa com uso diversificado. Cruz das Almas, BA: Embrapa Mandioca e Fruticultura Tropical, 2005.

GU, Z.; GU, L.; EILS, R.; SCHLESNER, M.; BRORS, B. Circlize implements and enhances circular visualization in R. **Bioinformatics**, v. 30, n. 19, p. 2811-2812, 2014. <https://doi.org/10.1093/bioinformatics/btu393>.

HOWELER, R.; LUTALADIO, N.; THOMAS, G. **Save and grow: cassava. A guide to sustainable production intensification**. FAO, 2013. <https://doi.org/10.1002/jsfa.2740310104>.

IBGE [INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA], 2019. **Levantamento Sistemático da Produção Agrícola**. Disponível em: <https://biblioteca.ibge.gov.br/visualizacao/periodicos/2415/epag_2019_dez.pdf>. Acesso em: 02 de Junho de 2021.

JARVIS, A.; RAMIREZ-VILLEGAS, J.; CAMPO, B. V. H.; NAVARRO-RACINES, C. Is cassava the answer to African climate change adaptation?. **Tropical Plant Biology**, v. 5, n. 1, p. 9-29, 2012. <https://doi.org/10.1007/s12042-012-9096-7>

JOMBART, T.; DEVILLARD, S.; BALLOUX, F. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. **BMC genetics**, v. 11, n. 1, p. 1-15, 2010. <https://doi.org/10.1186/1471-2156-11-94>.

KAWANO, K.; FUKUDA, W. M. G.; CENPUKDEE, U. Genetic and environmental effects on dry matter content of cassava root. **Crop Science**, v. 27, n. 1, p. 69-74, 1987. <https://doi.org/10.2135/cropsci1987.0011183X002700010018x>.

KOUNDINYA, A. V. V.; HEGDE, V.; SHEELA, M. N.; CHANDRA, C. V. Evaluation of cassava varieties for tolerance to water deficit stress conditions. **Journal of Root Crops**, v. 44, n. 1, p. 70-75, 2018.

MAYER, J. E.; PFEIFFER, W. H.; BEYER, P. Biofortified crops to alleviate micronutrient malnutrition. **Current Opinion in Plant Biology**, v. 11, n. 2, p. 166-170, 2008. <https://doi.org/10.1016/j.pbi.2008.01.007>.

MCKEY, D.; CAVAGNARO, T. R.; CLIFF, J.; GLEADOW, R. Chemical ecology in coupled human and natural systems: people, manioc, multitrophic interactions and global change. **Chemoecology**, v. 20, n. 2, p. 109-133, 2010. <http://dx.doi.org/10.1007/s00049-010-0047-1>.

MEENAKSHI, J. V.; JOHNSON, N. L.; MANYONG, V. M.; DEGROOTE, H.; JAVELOSA, J.; YANGGEN, D. R.; NAHER, F.; GONZALEZ, C.; GARCÍA, J.; MENG, E. How cost-effective is biofortification in combating micronutrient malnutrition? An ex ante assessment. **World Development**, v. 38, n. 1, p. 64-75, 2010. <https://doi.org/10.1016/j.worlddev.2009.03.014>

MEHARI, G.; AMSALU, N.; TEWODROS, M. Estimates of genetic components for yield and quality of cassava (*Manihot esculenta* Crantz) genotypes at Jimma, Southwest Ethiopia. **International Journal of Plant Breeding and Genetics**, v. 9, n. 1, p. 1-12, 2015. doi: 10.3923/ijpbg.2015.1.12.

MONTAGNAC, J. A.; DAVIS, C. R.; TANUMIHARDJO, S. A. Processing techniques to reduce toxicity and antinutrients of cassava for use as a staple food. **Comprehensive Reviews in Food Science and Food Safety**, v. 8, n. 1, p. 17-27, 2009. <https://doi.org/10.1111/j.1541-4337.2008.00064.x>.

MULAMBA, N.N.; MOCK, J.J. Improvement of yield potential of the Eto Blanco maize (*Zea mays* L.) population by breeding for plant traits. **Egyptian Journal of Genetics and Cytology**, v.7, p.40-57, 1978.

NHASSICO, D.; MUQUINGUE, H.; CLIFF, J.; CUMBANA, A.; BRADBURY, J. H. Rising African cassava production, diseases due to high cyanide intake and control measures. **Journal of the Science of Food and Agriculture**, v. 88, n. 12, p. 2043-2049, 2008. <https://doi.org/10.1002/jsfa.3337>.

NJOKU, D. N.; EGESI, C. N.; GRACEN, V. E.; OFFEI, S. K.; ASANTE, I. K.; DANQUAH, E. Y. Identification of pro-vitamin A cassava (*Manihot esculenta* Crantz) varieties for adaptation and adoption through participatory research. **Journal of Crop Improvement**, v. 28, n. 3, p. 361-376, 2014. <https://doi.org/10.1080/15427528.2014.888694>.

NJOKU, D. N.; GRACEN, V. E.; OFFEI, S. K.; ASANTE, I. K.; EGESI, C. N.; KULAKOW, P.; CEBALLOS, H. Parent-offspring regression analysis for total carotenoids and some agronomic traits in cassava. **Euphytica**, v. 206, n. 3, p. 657-666, 2015. <https://doi.org/10.1007/s10681-015-1482-4>.

NJUKWE, E.; HANNA, R., KIRSCHT, H., ARAKI, S. Farmers perception and criteria for cassava variety preference in Cameroon. **African Study Monographs**, v. 34, n. 4, p. 221-234, 2013. <https://doi.org/10.14989/185091>.

OLIVEIRA, E.J.; FERREIRA, C. F.; SANTOS, V.S.; JESUS, O. N.; OLIVEIRA, G. A. F.; SILVA, M.S. Potential of SNP markers for the characterization of Brazilian cassava germplasm. **Theoretical and Applied Genetics**, v. 127, n. 6, p. 1423-1440, 2014. <https://doi.org/10.1007/s00122-014-2309-8>.

OLIVEIRA, E. J.; MORGANTE, C. V.; AIDAR, S.T.; CHAVES, A.R.M.; ANTONIO, R. P.; CRUZ, J. L.; FILHO, M. A. C. Evaluation of cassava germplasm for drought tolerance under field conditions. **Euphytica**, v. 213, n. 8, p. 188, 2017. <https://doi.org/10.1007/s10681-017-1972-7>.

OLIVEIRA, E. J.; OLIVEIRA FILHO, O. S.; SANTOS, V. S. Classification of cassava genotypes based on qualitative and quantitative data. **Genetics and**

Molecular Research, v. 14, n. 1, p. 906-924, 2015.
<http://dx.doi.org/10.4238/2015>

OLIVEIRA, E. J.; SANTANA, F. A.; OLIVEIRA, L. A. D.; SANTOS, V. D. S. Genotypic variation of traits related to quality of cassava roots using affinity propagation algorithm. **Scientia Agricola**, v. 72, n. 1, p. 53-61, 2015. <https://doi.org/10.1590/0103-9016-2014-0043>.

PARKES, E.; AINA, O.; KINGSLEY, A.; ILUEBBEY, P.; BAKARE, M.; AGBONA, A.; AKPOTUZOR, P.; LABUSCHAGNE, M.; KULAKOW, P. Combining ability and genetic components of yield characteristics, dry matter content, and total carotenoids in provitamin A cassava F1 cross-progeny. **Agronomy**, v. 10, n. 12, p. 1850, 2020. <https://doi.org/10.3390/agronomy10121850>.

PARMAR, A.; STURM, B.; HENSEL, O. Crops that feed the world: Production and improvement of cassava for food, feed, and industrial uses. **Food Security**, v. 9, n. 5, p. 907-927, 2017. <https://doi.org/10.1007/s12571-017-0717-8>.

PATIL, I. **Visualizations with statistical details: The 'ggstatsplot' approach**. PsyArxiv, 2018. doi:10.31234/osf.io/p7mku.

PEPRAH, B. B.; PARKES, E. Y.; HARRISON, O. A.; VAN BILJON, A.; STEINER-ASIEDU, M.; LABUSCHAGNE, M. T. Proximate composition, cyanide content, and carotenoid retention after boiling of provitamin A-rich cassava grown in Ghana. **Foods**, v. 9, n. 12, p. 1800, 2020. <https://doi.org/10.3390/foods9121800>.

R CORE TEAM. 2021. **R: A language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna.

RABBI, I. Y.; UDOH, L. I.; WOLFE, M.; PARKES, E. Y.; GEDIL, M. A.; DIXON, A.; RAMU, P.; JANNINK, J-L.; KULAKOW, P. Genome-wide association

mapping of correlated traits in cassava: dry matter and total carotenoid content. **The plant genome**, v. 10, n. 3, 2017. <https://doi.org/10.3835/plantgenome2016.09.0094>.

RODRIGUEZ-AMAYA, D. B.; KIMURA, M. **Harvest Plus handbook for carotenoid analysis**. Washington: IFPRI, 2004. 58p.

SALTZMAN, A.; BIROL, E.; OPARINDE, A.; ANDERSSON, M. S.; ASARE-MARFO, D.; DIRESSIE, M. T.; GONZALEZ, C.; LIVIDINI, K.; MOURSI, M.; ZELLER, M. Availability, production, and consumption of crops biofortified by plant breeding: current evidence and future potential. **Annals of the New York Academy of Sciences**, v. 1390, n. 1, p. 104-114, 2017. doi: 10.1111/nyas.13314.

SÁNCHEZ, T.; CEBALLOS, H.; DUFOUR, D.; ORTIZ, D.; MORANTE, N.; CALLE, F.; FELDE, T. Z.; DOMÍNGUEZ, M.; DAVRIEUX, F. Prediction of carotenoids, cyanide and dry matter contents in fresh cassava root using NIRS and Hunter color techniques. **Food Chemistry**, v. 151, p. 444–451, 2014. <https://doi.org/10.1016/j.foodchem.2013.11.081>.

SÁNCHEZ, T.; CHÁVEZ, A. L.; CEBALLOS, H.; RODRIGUEZ-AMAYA, D. B.; NESTEL, P.; ISHITANI, M. Reduction or delay of post-harvest physiological deterioration in cassava roots with higher carotenoid content. **Journal of the Science of Food and Agriculture**, v. 86, n. 4, p. 634-639, 2006. <https://doi.org/10.1002/jsfa.2371>.

SAYRE, R.; BEECHING, J. R.; CAHOON, E. B.; EGESI, C.; FAUQUET, C.; FELLMAN, J.; FREGENE, M.; GRUISSEM, W.; MALLOWA, S.; MANARY, M.; MAZIYA-DIXON, B.; MBANASO, A.; SCHACHTMAN, D. P.; SIRITUNGA, D.; TAYLOR, N.; VANDERSCHUREN, H.; ZHANG, P. The BioCassava plus program: biofortification of cassava for sub-Saharan Africa. **Annual review of plant biology**, v. 62, p. 251-272, 2011. <https://doi.org/10.1146/annurev-arplant-042110-103751>.

SHETE, V.; QUADRO, L. Mammalian metabolism of β -carotene: Gaps in knowledge. **Nutrients**, v. 5, n. 12, p. 4849–4868, 2013. <https://doi.org/10.3390/nu5124849>.

SOUZA, L. D.; SOUZA, L. S.; GOMES, J. C. (2006) Exigências edáficas da cultura da mandioca. In: SOUZA, L. S.; FARIAS, A. R. N.; MATTOS, P. L. P.; FUKUDA, W. M. G. (Eds.), **Aspectos Socioeconômicos e Agronômicos da Mandioca**. Cruz Das Almas: Embrapa Mandioca e Fruticultura Tropical, p. 70–214.

STAPLETON, G. Global starch market outlook and competing starch raw materials for starches by product segment and region. **Cassava Starch World 2012**, v. 2012, p. 22-24, 2012.

TALEON, V.; SUMBU, D.; MUZHINGI, T.; BIDIKA, S. Carotenoids retention in biofortified yellow cassava processed with traditional African methods. **Journal of the Science of Food and Agriculture**, v. 99, n. 3, p. 1434-1441, 2019. doi.org/10.1002/jsfa.9347.

TALSMA, ELISE F.; BROUWER, I. D.; VERHOEF, H.; MBERA, G. N.; MWANGI, A. M.; DEMIR, A. Y.; MAZIYA-DIXON, B.; BOY, E.; ZIMMERMANN, M. B.; MELSE-BOONSTRA, A. Biofortified yellow cassava and vitamin A status of Kenyan children: a randomized controlled trial. **The American journal of clinical nutrition**, v. 103, n. 1, p. 258-267, 2016. <https://doi.org/10.3945/ajcn.114.100164>.

VASCONCELOS, L.M.; BRITO, A.C.; CARMO, C.D.; OLIVEIRA, P.H.; OLIVEIRA, E.J. Phenotypic diversity of starch granules in cassava germplasm. **Genetics and Molecular Research**, v.16(2), gmr16029276, 2017.

VIEIRA, E. A.; FIALHO, J. D. F.; FALEIRO, F. G.; BELLON, G.; FONSECA, K. G. D.; CARVALHO, L. J. C. B.; DENKE, M. L. Characterization of sweet cassava

accessions based on molecular, quantitative and qualitative data. **Crop Breeding and Applied Biotechnology**, v. 11, n. 3, p. 232-240, 2011. <https://doi.org/10.1590/S1984-70332011000300005>.

WEI T, SIMKO V. **R package "corrplot": Visualization of a Correlation Matrix**. (Version 0.88), 2021. Disponível em: <<https://github.com/taiyun/corrplot>>. Acesso em: 05 de junho de 2021.

WELSCH, R.; ARANGO, J.; BÄR, C.; SALAZAR, B.; AL-BABILI, S.; BELTRÁN, J.; BEYER, P. Provitamin A accumulation in cassava (*Manihot esculenta*) roots driven by a single nucleotide polymorphism in a phytoene synthase gene. **Plant Cell**, v. 22, p. 3348–3356, 2010. <https://doi.org/10.1105/tpc.110.077560>.

Tabela Suplementar 1. Acessos de mandioca avaliados para teor de carotenoides e atributos produtivos e de qualidade de raízes.

Clone	Country	State	City	Region
Aciolina	Brazil	Ceará	Redenção	Landrace
Aipim-Manteiga	Brazil	Bahia	Juazeiro	Landrace
BGM-0019	Brazil	Pará	Bragança	Landrace
BGM-0027	Brazil	-	-	Landrace
BGM-0039	Brazil	Bahia	Cruz das Almas	Landrace
BGM-0054	Brazil	Pará	Bragança	Landrace
BGM-0058	Brazil	-	-	Landrace
BGM-0060	Brazil	-	-	Landrace
BGM-0061	Brazil	Pará	-	Landrace
BGM-0067	Brazil	Pará	-	Landrace
BGM-0069	Brazil	Bahia	Boa Vista do Tupim	Landrace
BGM-0070	Brazil	Pará	Bragança	Landrace
BGM-0120	Brazil	Bahia	Ilha dos Frades	Landrace
BGM-0122	Brazil	Pará	Bragança	Landrace
BGM-0123	Brazil	Bahia	Cruz das Almas	Landrace
BGM-0131	Brazil	Amazonas	-	Landrace
BGM-0132	Brazil	Amazonas	-	Landrace
BGM-0162-B	Brazil	Bahia	Cruz das Almas	Landrace
BGM-0212	Brazil	Bahia	Boa Vista do Tupim	Landrace
BGM-0239	Brazil	-	-	Landrace
BGM-0257	Brazil	Bahia	Itaparica	Landrace
BGM-0267	Brazil	São Paulo	Campinas	Improved variety
BGM-0269	Brazil	Sergipe	Araúá	Landrace
BGM-0285	Brazil	-	-	Landrace
BGM-0290	Colombia	ValledelCauca	-	Improved variety
BGM-0327	Colombia	ValledelCauca	Cali	Improved variety
BGM-0330	Colombia	Valle	Cali	Improved Variety
BGM-0345	Colombia	ValledelCauca	Cali	Improved variety
BGM-0364	Brazil	Ceará	Pentecoste	Landrace
BGM-0368	Brazil	Ceará	Paraipaba	Landrace
BGM-0380	Brazil	Ceará	Pedra Branca	Landrace
BGM-0408	Brazil	-	-	Landrace
BGM-0424	Brazil	-	-	Landrace
BGM-0444	Brazil	São Paulo	Campinas	Improved variety
BGM-0468	Brazil	Rio de Janeiro	Itaguaí	Improved variety
BGM-0470	Brazil	-	-	Landrace
BGM-0472	Brazil	Rio de Janeiro	Seropédica	Improved variety
BGM-0507	Brazil	Mato Grosso	-	Landrace
BGM-0579	Brazil	Pernambuco	Itambé	Landrace
BGM-0597	Brazil	Rio Grande do Sul	São Sebastião do Caí	Landrace
BGM-0614	Brazil	São Paulo	Campinas	Improved variety
BGM-0659	Brazil	Maranhão	Vitorino Freire	Landrace
BGM-0660	Brazil	Maranhão	Monção	Landrace
BGM-0661	Brazil	Rio Grande do Sul	Santa Rosa	Landrace
BGM-0672	Brazil	Maranhão	Coroatá	Landrace
BGM-0673	Brazil	Maranhão	São Domingos do Maranhão	Landrace
BGM-0675	Brazil	Maranhão	Santa Inês	Landrace
BGM-0677	Brazil	Maranhão	Pindari Mirim	Landrace
BGM-0678	Brazil	Maranhão	Coroatá	Landrace
BGM-0679	Brazil	Maranhão	Monção	Landrace
BGM-0685	Brazil	Maranhão	Coroatá	Landrace
BGM-0694	Brazil	Maranhão	Codó	Landrace
BGM-0699	Brazil	Maranhão	Codó	Landrace
BGM-0728	Brazil	Bahia	Cruz das Almas	Improved variety
BGM-0729	Brazil	Bahia	Cruz das Almas	Improved variety
BGM-0733	Brazil	Espírito Santo	Linhares	Landrace

BGM-0760	Brazil	Espírito Santo	Boa Esperança	Landrace
BGM-0776	Brazil	Bahia	Alcobaça	Landrace
BGM-0779	Brazil	Bahia	Alcobaça	Landrace
BGM-0780(B)	-	-	-	-
BGM-0793	Brazil	Bahia	Eunápolis	Landrace
BGM-0815	Brazil	Alagoas	Arapiraca	Landrace
BGM-0870	Colombia	Valle	Cali	Improved Variety
BGM-0872	Brazil	Amapá	Oiapoque	Landrace
BGM-0878	Brazil	Amazonas	-	Landrace
BGM-0879	Brazil	Santa Catarina	navegantes	Landrace
BGM-0882	Colombia	ValledelCauca	Cali	Improved variety
BGM-0884	Colombia	ValledelCauca	Cali	Improved variety
BGM-0886	Brazil	Amazonas	Lábrea	Landrace
BGM-0889	Brazil	Tocantins	Tocantinópolis	Landrace
BGM-0893	Brazil	Amazonas	Lábrea	Landrace
BGM-0895	Brazil	Pará	Belém	Landrace
BGM-0901	Brazil	Ceará	-	Landrace
BGM-0902	Brazil	Pará	Bragança	Landrace
BGM-0909	Brazil	Pará	Bragança	Landrace
BGM-0912	Brazil	-	-	Landrace
BGM-0914	Brazil	Pará	Curralinho	Landrace
BGM-0918	Brazil	Pará	Cametá	Landrace
BGM-0919	Brazil	Amazonas	-	Landrace
BGM-0920	Brazil	Amazonas	-	Landrace
BGM-0936	Brazil	Bahia	Vitória da Conquista	Landrace
BGM-0940	Brazil	Pará	Bragança	Landrace
BGM-0941	Brazil	Amazonas	Tapauá	Landrace
BGM-0942	Brazil	Amazonas	Tapauá	Landrace
BGM-0943	Brazil	Amazonas	-	Landrace
BGM-0945	Brazil	Amazonas	Manacapuru	Landrace
BGM-0951	Brazil	Amazonas	Manaus	Landrace
BGM-0952	Brazil	Amazonas	Caracarái	Landrace
BGM-0953	Brazil	Amazonas	Aripuanã	Landrace
BGM-0954	Brazil	Amazonas	Itacoatiara	Landrace
BGM-0958	Brazil	Amazonas	Parintins	Landrace
BGM-0961	Brazil	Amazonas	Tefé	Landrace
BGM-0962	Brazil	Amazonas	Careiro	Landrace
BGM-0963	Brazil	Amazonas	Itacoatiara	Landrace
BGM-0966	Brazil	Amazonas	Barreirinha	Landrace
BGM-0967	Brazil	Amazonas	Urucará	Landrace
BGM-0968	Brazil	Amazonas	Manaus	Landrace
BGM-0971	Brazil	Amazonas	-	Landrace
BGM-0974	Brazil	Amazonas	Parintins	Landrace
BGM-0975	Brazil	Amazonas	Urucará	Landrace
BGM-0976	Brazil	Amazonas	Manaus	Landrace
BGM-0983	Brazil	Amazonas	-	Landrace
BGM-0986	Brazil	Amazonas	-	Landrace
BGM-0987	Brazil	Amazonas	-	Landrace
BGM-0991	Brazil	Amazonas	Parintins	Landrace
BGM-0993	Brazil	Amazonas	Irاندوبا	Landrace
BGM-0994	Brazil	Amazonas	Rio Preto da Eva	Landrace
BGM-0995	Colombia	ValledelCauca	Cali	Improved variety
BGM-0996	Colombia	ValledelCauca	Cali	Improved variety
BGM-1034	Brazil	Goiás	Taguatinga	Landrace
BGM-1043	Brazil	Piauí	Cristino Castro	Landrace
BGM-1133	Brazil	Amazonas	Autazes	Landrace
BGM-1136	Brazil	Amazonas	Paraná de Paracuúba	Landrace
BGM-1137	Brazil	Amazonas	Careiro	Landrace
BGM-1141	Brazil	Amazonas	-	Landrace

BGM-1142	Brazil	Amazonas	-	Landrace
BGM-1143	Brazil	Mato Grosso	Aripuanã	Landrace
BGM-1144	Brazil	Amazonas	Barreirinha	Landrace
BGM-1146	Brazil	Amazonas	Barreirinha	Landrace
BGM-1147	Brazil	Amazonas	-	Landrace
BGM-1152	Brazil	Amazonas	-	Landrace
BGM-1159	Brazil	-	-	Landrace
BGM-1162	Brazil	-	-	Landrace
BGM-1164	Brazil	-	-	Landrace
BGM-1168	Brazil	-	-	Landrace
BGM-1177	Brazil	-	-	Landrace
BGM-1178	Brazil	-	-	Landrace
BGM-1190	Brazil	Pará	Alenquer	Landrace
BGM-1191	Brazil	Pará	Anajas	Landrace
BGM-1198	Brazil	Amapá	Mazagão	Landrace
BGM-1210	Brazil	Pará	Bragança	Landrace
BGM-1213	Colombia	ValledelCauca	Calí	Improved variety
BGM-1214	Brazil	Pará	São Domingos do Capim	Landrace
BGM-1224	Brazil	-	-	Landrace
BGM-1226	Brazil	-	-	Landrace
BGM-1227	Brazil	-	-	Landrace
BGM-1280	Brazil	Pernambuco	Ouricuri	Landrace
BGM-1284	Brazil	Pernambuco	Ouricuri	Landrace
BGM-1286	Brazil	Pernambuco	Araripina	Landrace
BGM-1296	Brazil	Pernambuco	Araripina	Landrace
BGM-1304	Brazil	Pernambuco	Exu	Landrace
BGM-1310	Brazil	Ceará	Crato	Landrace
BGM-1339	Brazil	Pernambuco	Sertânia	Landrace
BGM-1355	Brazil	Paraíba	Teixeira	Landrace
BGM-1390	Brazil	Rio Grande do Norte	Jaçanã	Landrace
BGM-1496	Brazil	Bahia	Itiúba	Landrace
BGM-1497	Brazil	Bahia	Itiúba	Landrace
BGM-1498	Brazil	Bahia	Itiúba	Landrace
BGM-1518	Brazil	Bahia	Cafarnaum	Landrace
BGM-1585	Brazil	Piauí	Itaueira	Landrace
BGM-1586	Brazil	Piauí	Itaueira	Landrace
BGM-1588	Brazil	Piauí	Floriano	Landrace
BGM-1594	Brazil	Piauí	São Gonçalo do Piauí	Landrace
BGM-1596	Brazil	Piauí	Barro Duro BR316	Landrace
BGM-1605	Brazil	Piauí	Oeiras	Landrace
BGM-1610	Brazil	Piauí	Simplicio Mendes	Landrace
BGM-1614	Brazil	Piauí	Paulistana	Landrace
BGM-1659	Brazil	Bahia	Cruz das Almas	Improved variety
BGM-1666	Brazil	Pará	BR316	Landrace
BGM-1674	Brazil	Pará	BR316	Landrace
BGM-1675	Brazil	Pará	BR316	Landrace
BGM-1682	Brazil	Bahia	Tucano	Landrace
BGM-1690	Brazil	Pernambuco	Araripina	Landrace
BGM-1692	Brazil	Bahia	São Felipe	Landrace
BGM-1700	Brazil	Amazonas	BR230	Improved variety
BGM-1701	Brazil	Amazonas	BR230	Improved variety
BGM-1702	Brazil	Amazonas	BR230	Improved variety
BGM-1703	Brazil	Amazonas	BR230	Improved variety
BGM-1704	Brazil	Amazonas	BR230	Improved variety
BGM-1705	Brazil	Amazonas	BR230	Improved variety
BGM-1706	Brazil	Amazonas	BR230	Improved variety
BGM-1708	Brazil	Amazonas	BR230	Improved variety
BGM-1709	Brazil	Amazonas	BR230	Improved variety
BGM-1710	Brazil	Amazonas	BR230	Improved variety

BGM-1712	Brazil	Amazonas	BR174	Improved variety
BGM-1716	Brazil	Amazonas	BR178	Improved variety
BGM-1718	Brazil	Amazonas	BR179	Improved variety
BGM-1721	Brazil	Bahia	Santo Antônio de Jesus	Landrace
BGM-1722-D	-	-	-	-
BGM-1725	Brazil	Bahia	Cruz das Almas	Improved variety
BGM-1732	Brazil	Maranhão	BR226	Landrace
BGM-1734	Brazil	Maranhão	Santa Rita	Landrace
BGM-1735	Brazil	Maranhão	Santa Rita	Landrace
BGM-1737	Brazil	Maranhão	Santa Rita	Landrace
BGM-1739	Brazil	Maranhão	Santa Rita	Landrace
BGM-1741	Brazil	Maranhão	Santa Rita	Landrace
BGM-1742	Brazil	Maranhão	Santa Rita	Landrace
BGM-1743	Brazil	Maranhão	Santa Rita	Landrace
BGM-1744	Brazil	Maranhão	Santa Rita	Landrace
BGM-1745	Brazil	Maranhão	Santa Rita	Landrace
BGM-1746	Brazil	Maranhão	Santa Rita	Landrace
BGM-1747	Brazil	Maranhão	Santa Rita	Landrace
BGM-1752	Brazil	Maranhão	Chapadinha	Landrace
BGM-1754	Brazil	Maranhão	Chapadinha	Landrace
BGM-1756	Brazil	Maranhão	Chapadinha	Landrace
BGM-1759	Brazil	Maranhão	Chapadinha	Landrace
BGM-1764	Brazil	Maranhão	Chapadinha	Landrace
BGM-1766	Brazil	Maranhão	Chapadinha	Landrace
BGM-1768	Brazil	Maranhão	Chapadinha	Landrace
BGM-1773	Brazil	Maranhão	Chapadinha	Landrace
BGM-1779	Brazil	Maranhão	São Luís	Landrace
BGM-1780	Brazil	Maranhão	São Luís	Landrace
BGM-1783	Brazil	Maranhão	São Luís	Landrace
BGM-1784	Brazil	Maranhão	São Luís	Landrace
BGM-1814	Brazil	Bahia	Vitória da Conquista	Landrace
BGM-1818	Brazil	Bahia	Caetité	Landrace
BGM-1828	Brazil	Pará	Belém	Improved variety
BGM-1832	Brazil	Pará	Belém	Improved variety
BGM-1834	Brazil	Pará	Belém	Improved variety
BGM-1835	Brazil	Pará	Belém	Improved variety
BGM-1866	Brazil	Paraná	Br277	Landrace
BGM-1873	Brazil	Mato Grosso	Primavera do Leste	Landrace
BGM-1878	Brazil	Tocantins	BR277	Landrace
BGM-1884	Brazil	Amazonas	Parintins	Landrace
BGM-1956	Brazil	Amazonas	Nhamundá	Landrace
BGM-1957	Brazil	Amazonas	Juruti	Landrace
BGM-1962	Brazil	Amazonas	Juruti	Landrace
BGM-1969	Brazil	Amazonas	Juruti	Landrace
BGM-1980	Brazil	Amazonas	Itapiranga	Landrace
BGM-1985	Brazil	Amazonas	Itapiranga	Landrace
BGM-2024	Brazil	Paraná	Marechal Cândido Rondon	Improved variety
BGM-2029	Brazil	Pernambuco	Araripina	Landrace
BGM-2041	Brazil	Bahia	Cruz das Almas	Improved variety
BGM-2047	Colombia	ValledelCauca	Cali	Improved variety
BGM-2063	Brazil	Rio Grande do Sul	Pelotas	Improved variety
BGM-2074	Brazil	Rio Grande do Sul	Pelotas	Improved variety
BGM-2097	Brazil	Paraná	Marechal Cândido Rondon	Landrace
BGM-2119	Brazil	Paraná	Marechal Cândido Rondon	Improved variety
BGM-2123	Brazil	Bahia	Petrolina	Improved variety
BGM-2127	Brazil	Distrito Federal	Planaltina	Landrace
BGM-2156	Brazil	Pernambuco	Petrolina	Improved variety
BGM-2167	Brazil	Espírito Santo	lúna	Landrace
BGM-2187	Brazil	Pernambuco	Petrolina	Landrace

BGM-2191	Brazil	Pernambuco	Petrolina	Landrace
BGM-2210	Brazil	Pernambuco	Petrolina	Landrace
BGM-2212	Brazil	Pernambuco	Petrolina	Landrace
BGM-2293	Brazil	Pernambuco	Petrolina	Landrace
BGM-2342	Colombia	ValledelCauca	Cali	Improved variety
BGM-2343	Colombia	ValledelCauca	Cali	Improved variety
BGM-2344	Colombia	ValledelCauca	Cali	Improved variety
BGM-2345	Colombia	ValledelCauca	Cali	Improved variety
BGM-2346	Colombia	ValledelCauca	Cali	Improved variety
BGM-2348	Colombia	ValledelCauca	Cali	Improved variety
BGM-2349	Brazil	Pernambuco	Petrolina	Landrace
BGM-2353	Brazil	Bahia	Cruz das Almas	Landrace
BGM-2357	Brazil	Bahia	Cruz das Almas	Landrace
BRS-396	Brazil	Distrito Federal	Brasília	Improved variety
BRS-399	Brazil	Distrito Federal	Brasília	Improved variety
BRS-Dourada	Brazil	Bahia	Cruz das Almas	Improved variety
BRS-Gema-de-Ovo	Brazil	Bahia	Cruz das Almas	Improved variety
Conquista-1	Brazil	Roraima	Boa Vista	Landrace
CPAFRO-03	Brazil	Rondonia	Ariquemes	Landrace
CPAFRO-04	Brazil	-	-	Landrace
CPAFRO-06	Brazil	-	-	Landrace
CPAFRO-08	Brazil	-	-	Landrace
CPAFRO-09	Brazil	Acre	Rio Branco	Landrace
CPAFRO-43	Brazil	Paraíba	Campina Grande	Landrace
Inajazinha	Brazil	Roraima	Boa Vista	Landrace
Jacona	Brazil	-	-	-
Joao-Florestal-5	Brazil	Minas Gerais	Florestal	Landrace
Joao-Florestal-6	Brazil	Minas Gerais	Florestal	Landrace
Pioneira	Brazil	Pará	-	Landrace
Roxinha	Brazil	-	-	-
Sao-Francisco-de-Paula	Brazil	Minas Gerais	São Francisco de Paula	Landrace
Uberlandia-2	Brazil	Minas Gerais	Uberlandia	Landrace

ARTIGO 2

FENOTIPAGEM BASEADA EM IMAGENS DIGITAIS DE RAÍZES DE MANDIOCA PARA ESTUDOS DE DIVERSIDADE E PREDIÇÃO DO TEOR DE CAROTENOIDES

FENOTIPAGEM BASEADA EM IMAGENS DIGITAIS DE RAÍZES DE MANDIOCA PARA ESTUDOS DE DIVERSIDADE E PREDIÇÃO DO TEOR DE CAROTENOIDES

RESUMO: A fenotipagem para quantificação do teor total de carotenoides é bastante sensível, demorada, tediosa e de alto custo. O desenvolvimento de ferramentas de fenotipagem rápida em larga escala são fundamentais para a triagem de centenas de genótipos de mandioca em um curto período de tempo no programa de biofortificação. Este estudo objetivou: (i) usar imagens digitais para extrair informações sobre a cor da polpa de raízes e estimar suas correlações com o teor de carotenoides totais (TCC); (ii) selecionar modelos para predição do TCC com base nos índices colorimétricos. Imagens *Red Green Blue* (RGB) foram capturadas em amostras de raízes de 228 genótipos biofortificados e a diferença na coloração foi analisada por meio dos índices L^* , a^* , b^* , *Hue* e *croma* do sistema de cores CIELAB (*International Commission on Illumination*) e luminosidade. Os dados colorimétricos foram utilizados para análise de componentes principais (PCA), correlação e construção de modelos de predição do TCC com base em regressão e aprendizado de máquina. Alta correlação positiva entre TCC e as variáveis b^* ($r= 0,90$) e *croma* ($r=0,89$) foi identificada, enquanto que as demais correlações foram medianas e negativas, e o parâmetro L^* não apresentou correlação significativa com TCC. De modo geral a acurácia da maioria dos modelos de predição (com todas as variáveis e apenas com as mais importantes) foi elevada (R^2 variando de 0,81 a 0,94). Entretanto, o modelo de predição *artificial neural network* apresentou a melhor capacidade preditiva ($R^2=0,94$), associada ao menor erro nas estimativas do TCC (RMSE= 0,24). A estrutura da população estudada revelou cinco grupos e alta variabilidade genética com base na PCA quanto aos índices colorimétricos e conteúdo de carotenoides totais. Nossos resultados demonstraram que o uso de dados obtidos a partir da análise de imagens digitais é uma alternativa econômica, rápida e eficaz para o desenvolvimento de ferramentas de fenotipagem do TCC em raízes de mandioca com alta capacidade preditiva.

Palavras-chave: Visão computacional, CIELAB, modelos preditivos.

IMAGE-BASED PHENOTYPING OF CASSAVA ROOTS FOR DIVERSITY STUDIES AND CAROTENOIDS PREDICTION

ABSTRACT: Phenotyping to quantify the total content of carotenoids is very sensitive, time-consuming, tedious and costly. The development of large-scale rapid phenotyping tools is essential for screening hundreds of cassava genotypes in a short period of time in the biofortification program. This study aimed: (i) use digital images to extract information on the pulp color of cassava roots and estimate correlations with total carotenoid content (TCC); (ii) select predictive models for TCC with high accuracy using colorimetric indices. Red Green Blue (RGB) images were captured in root samples from 228 biofortified genotypes and the difference in color was analyzed using L^* , a^* , b^* , Hue and chroma indices from the CIELAB (International Commission on Illumination) color system and luminosity. Colorimetric data were used for principal component analysis (PCA), correlation and for developing prediction models for TCC based on regression and machine learning. Colorimetric data were used for PCA, correlation and TCC prediction models based on regression and machine learning. A high positive correlation between TCC and the variables b^* ($r=0.90$) and chroma ($r=0.89$) was identified, while the other correlations were median and negative, and the L^* parameter did not present a significant correlation with TCC. In general, the accuracy of most prediction models (with all variables and only the most important ones) was high (R^2 ranging from 0.81 to 0.94). However, the artificial neural network prediction model presented the best predictive capacity ($R^2=0.94$), associated with the smallest error in the TCC estimates (RMSE=0.24). The structure of the studied population revealed five groups and high genetic variability based on PCA regarding colorimetric indices and total carotenoid content. Our results demonstrated that the use of data obtained from digital image analysis is an economical, fast and effective alternative for the development of TCC phenotyping tools in cassava roots with high predictive capacity.

Keywords: Computer vision, CIELAB, predictions models.

1. INTRODUÇÃO

A mandioca (*Manihot esculenta* Crantz) pertencente à família Euphorbiaceae é uma espécie amplamente cultivada em praticamente todo o mundo, por apresentar características bastante desejáveis para diversas espécies comerciais como a tolerância ao déficit hídrico, com ampla adaptação a diversas condições edafoclimáticas e produtividades satisfatórias, mesmo em solos com baixa fertilidade (BURNS et al., 2010; JARVIS et al., 2012). A Nigéria com 59,19 milhões de toneladas cultivados em 7,21 milhões de hectares é o maior produtor mundial, enquanto o Brasil ocupa a quinta colocação com produção de 17,49 milhões de toneladas em uma área de 1,19 milhões de hectares, e rendimento médio de 14,70 t ha⁻¹ (FAO, 2020).

As raízes de mandioca são compostas basicamente por carboidratos, o que a torna uma importante fonte de energia para milhões de pessoas em todo o mundo (FERRARO et al., 2016). As raízes ainda possuem em sua composição glicosídeos cianogênicos (linamarina e lotaustralina) que variam de 9,90 a 3927,0 µg.g⁻¹ (SÁNCHEZ et al., 2014; PEPRAH et al., 2020; PARMAR; STURM; HENSEL, 2017) e baixa qualidade nutricional, com valores consideravelmente reduzidos de proteínas, minerais (Zn, Fe, N, Ca, P, K e Mg) e vitaminas (B1, B2, B3, C e β-caroteno) (SAYRE et al., 2011; PARMAR; STURM; HENSEL, 2017).

O incremento do conteúdo nutricional por meio da biofortificação é uma estratégia que vem agregando valor nutricional às variedades de mandioca (SALTZMAN et al., 2017). A biofortificação é uma iniciativa internacional e interdisciplinar que busca reduzir a desnutrição humana aumentando a concentração de micronutrientes em diversas culturas básicas, à exemplo do aumento do teor de carotenoides em mandioca (CEBALLOS et al., 2017; SALTZMAN et al., 2017). Em raízes de mandioca busca-se o aumento do conteúdo de β-caroteno (carotenoide pró vitamina A), com a principal finalidade de combater a fome oculta (*hidden hunger*) ocasionada pela deficiência de vitamina A (SALTZMAN et al., 2017; ALAMU et al., 2019). Além disso, os programas de biofortificação procuram ajustar os níveis de outros atributos como elevada performance agrônômica (FUHRMANN et al., 2019), baixos

teores de compostos cianogênicos e altos teores de matéria seca nas raízes, visando atender preferências do mercado de mandioca de mesa (SÁNCHEZ et al., 2014; SALTZMAN et al., 2017).

Apesar dos avanços importantes no aumento do teor de carotenoides obtidos em diversos programas de melhoramento genético convencional (CEBALLOS; HERSHEY; BECERRA-LÓPEZ-LAVALLE, 2012; CEBALLOS et al., 2013; SÁNCHEZ et al., 2014), a fenotipagem para quantificação do teor total de carotenoides ou de β -caroteno é bastante sensível, demorada, tediosa e de alto custo. Por isso, novos métodos de fenotipagem rápida são necessários, uma vez que os programas de biofortificação requerem a triagem de centenas de amostras em um curto período de tempo (SÁNCHEZ et al., 2014; CEBALLOS et al., 2017).

Recentemente, predições com base em NIRS (*near infrared spectroscopy*) têm demonstrado alto potencial para fenotipagem indireta do teor de carotenoides em mandioca (SÁNCHEZ et al., 2014; BELALCAZAR et al., 2016; IKEOGU et al., 2017; ALAMU et al., 2019; IKEOGU et al., 2019). Além disso, metodologias baseadas na análise de imagens digitais permitem extrair informações de cor (BAI et al., 2018) devido à forte correlação entre dados digitais e visuais (WALTER et al., 2019). Uma vez que os carotenoides presentes possuem forte correlação com a intensidade da cor amarela (CEBALLOS et al., 2013), é presumível que este tipo de fenotipagem seja uma opção viável e adequada para a quantificação do teor de carotenoides.

A fenotipagem de raízes com base em imagens possui como principais vantagens a economia no tempo de análise, uso de câmeras digitais disponíveis comercialmente, de fácil manuseio, transporte e uso de softwares de código aberto para o processamento das imagens, demonstrando alto potencial para uso em programas de melhoramento de plantas (CORTES et al., 2017; WALTER et al., 2019). As imagens gravadas ainda permitem aos pesquisadores voltar e reexaminá-las sempre que surgirem dúvidas quanto ao processo de fenotipagem (BAI et al., 2018). Além disso, uma vez calibrados os modelos de predição é possível reduzir o número de amostras laboratoriais, concentrando esforços apenas naquelas de maior interesse e assim, reduzir em muito o custo das análises necessárias à seleção dos genótipos.

A fenotipagem rápida por meio de imagens já vem sendo amplamente utilizada para estudar a relação entre índices de cor extraídos das imagens e teor de carotenoides em diferentes culturas como cenoura (*Daucus carota*) (PACE et al., 2013), tomate (*Solanum lycopersicum* L.) (STINCO et al., 2013), azeite de oliva (*Olea europaea*) (SANAEIFAR; JAFARI; GOLMAKANI, 2018) e tangerina (*Citrus reticulata* Blanco) (HADIMAN; MITTAL, 2019). Normalmente, as imagens são obtidas por um dispositivo digital e salvas no espaço de cores *Red Green Blue* (RGB) tridimensional; no entanto este espaço não é percentualmente uniforme e não representa cores percebidas naturalmente pela visão humana (PATHARE; OPARA; AL-SAID, 2013). Alguns estudos tem optado por converter a matrizes de cores RGB em códigos CIELAB (CIE, 1978), quando há intenção de correlacionar as cores das imagens com algum atributo específico em plantas (YE; IZAWA; ZHANG, 2018; DODIG et al., 2019; LI; FRANK; MIGICOVSKY, 2019). O modelo CIE (CIE, 1978) de cores é uma escolha predominante entre pesquisadores (PATHARE; OPARA; AL-SAID, 2013) porque dispõe de espaço de cores com técnica de medição padronizada (DARRIGUES et al., 2008). O espaço de cores CIELAB (CIE, 1978) demonstrou ser bastante acurado na estimativa do teor de carotenoides com base em imagens digitais em tomate (*Solanum lycopersicum*) e em pólen de espécies das famílias Brassicaceae, Myrtaceae e Fabaceae (YE; IZAWA; ZHANG, 2018; GONZÁLEZ et al., 2020).

Em outras espécies, alta capacidade de predição de características associadas à qualidade nutricional tem sido observada quando utilizado índices colorimétricos obtidos de imagens digitais. Por exemplo, em tangerina (*Citrus reticulata* Blanco), os modelos *partial least squares* ($R^2 = 0,96$) e regressão linear de terceira ordem ($R^2 = 0,95$) foram capazes de predizer a maturação dos frutos (*citrus color index*) com alta precisão (HADIMAN; MITTAL, 2019). Em azeite de oliva (*Olea europaea*), os modelos *support vector machine* ($R^2 = 0,96$) e redes neurais artificiais ($R^2 = 0,94$) também apresentaram alta capacidade preditiva do teor de carotenoides nesta espécie (SANAEIFAR; JAFARI; GOLMAKANI, 2018). Adicionalmente, um modelo de regressão linear múltipla foi capaz de predizer o teor de carotenoides com alta precisão ($R^2 = 0.89$) em

pólen de espécies das famílias Brassicaceae, Myrtaceae, Fabaceae (GONZÁLEZ et al., 2020).

Considerando a necessidade de desenvolvimento de metodologias de fenotipagem de alto rendimento para otimizar a seleção dos programas de melhoramento de mandioca de mesa biofortificada, o presente estudo teve como objetivos: (i) avaliar o potencial de uso de imagens digitais para captar informações da cor de polpa de raízes de mandioca e estimar as correlações com o teor de carotenoides totais; (ii) selecionar modelos preditivos alimentados por índices colorimétricos que possuam alta precisão e capacidade preditiva do teor de carotenoides totais.

2. MATERIAL E MÉTODOS

2.1 Material vegetal e delineamento experimental

O conjunto de calibração foi composto por 228 genótipos do Banco de germoplasma (BAG) da Embrapa Mandioca e Fruticultura, Cruz das Almas, Bahia, Brasil (Latitude: 12° 40' 12" S, Longitude: 39° 06' 07" W e Altitude: 220m). Este painel de germoplasma é composto por genótipos com alta diversidade para teor de carotenoides, compostos cianogênicos, teor de matéria seca nas raízes e diversos atributos produtivos. Os clones foram cultivados de junho de 2019 a julho de 2020 na Estação Experimental da Embrapa Mandioca e Fruticultura, utilizando delineamento em blocos aumentados, tendo 34 variedades locais e melhoradas como testemunhas comuns. A parcela experimental foi composta por duas linhas com 10 plantas cada (20 plantas por parcela), com espaçamento de 0,90 m entre linhas e 0,80 m entre plantas. O plantio foi realizado com manivas de 16 cm em condição de sequeiro, seguindo as recomendações e práticas agrícolas para a cultura (SOUZA; SOUZA; GOMES, 2006). Para compor o conjunto de calibração foram previamente selecionados clones com coloração de polpa de raiz creme e amarela.

O clima na região de cultivo do germoplasma é do tipo tropical quente e úmido, sem estação seca, com pluviosidade média anual de 1.170 mm, onde os meses de março a agosto se apresentam como os mais chuvosos e de

setembro a fevereiro, os mais secos. A temperatura média anual é de 24,5° C e umidade relativa de 80%. Os solos predominantes da região são Latossolos Amarelos Distróficos.

2.2 Obtenção e processamento das imagens

Uma caixa (24x18x19 cm) com as paredes internas brancas e com o fundo forrado com material plástico azul para aumentar a precisão da separação dos pixels da raiz foi utilizada como uma plataforma de fenotipagem. Para obtenção das imagens digitais foi utilizado um tablet com sistema Android 9.0 e câmera com 8MP, com auxílio do flash para uniformizar a luminosidade, acoplado à parte superior da caixa.

Durante a colheita foram selecionadas entre 4-6 raízes padrão de todos os 228 genótipos. As raízes foram lavadas com água corrente para remover o excesso de solo aderido e em seguida a região central foi cortada em pedaços de aproximadamente 2 cm de espessura. Seis pedaços de raízes por genótipo provenientes foram colocados no centro do campo de visão da câmera e três imagens RGB foram capturadas. As imagens foram salvas no formato Joint Photographic Experts Group (JPEG) com resolução de 2448 x 3264 pixels. Foi realizado um pré-processamento das imagens utilizando os filtros *Gaussian Blur* e *remove background* com auxílio do software *ImageJ* versão 1.52a. A análise colorimétrica foi realizada utilizando o software *Tomato Analyzer - Color Test* versão 4.0 (DARRIGUES et al., 2008). Os valores iniciais dos componentes cromáticos (L^* , a^* e b^*) do sistema CIELAB (CIE, 1978), foram obtidos utilizando a calibração padrão do software. Esta calibração foi realizada assumindo valores de 1,0 para o coeficiente angular e de zero para os interceptos de uma equação de regressão linear. O excedente das raízes foi processado com auxílio de um multiprocessador de alimentos e reservadas para análise de carotenoides totais.

2.3 Avaliação da cor das raízes

As diferenças na coloração das raízes entre os diferentes genótipos de mandioca foram avaliadas por meio dos parâmetros CIELAB (CIE, 1978), os quais foram obtidos a partir dos valores RGB. A coordenada L^* indica escuridão

(0) para claridade (100) da cor, enquanto a^* e b^* indicam a variação do verde ao vermelho (-60 a 60) e do azul ao amarelo (-60 a 60), respectivamente. Para cada amostra, cada registro representa uma média de seis medições. As coordenadas cilíndricas *croma* e *Hue* foram calculadas a partir de a^* e b^* . O *Hue* (matiz) é uma medida angular de 0° a 360°, que representa a cor básica. Seu valor foi calculado com base nas seguintes equações: $H^\circ = \tan^{-1} \left(\frac{b^*}{a^*} \right)$, quando $a^* > 0$ e $b^* \geq 0$, e $H^\circ = 180 + \tan^{-1} \left(\frac{b^*}{a^*} \right)$, quando $a^* < 0$. O valor de *croma* é a média da saturação representada pelas médias de a^* e b^* .

Também foi calculada a luminosidade, definida pela média da quantidade de luminosidade (variando de 0 a 240) em todos os pixels. A luminosidade é responsável pela sensibilidade variável do olho humano à radiação em vários comprimentos de onda e define o brilho. A luminosidade foi estimada a partir do valor de RGB de cada pixel por meio da expressão:

$$Luminosidade = \frac{[\text{máximo}(R,G,B) + \text{mínimo}(R,G,B) \times 240]}{2 \times 255}$$

2.4 Quantificação do teor de carotenoides totais

Para a análise do teor carotenoides totais (TCC), foram coletadas duas amostras contendo 10 g, 15 g ou 25 g das raízes trituradas (a depender da intensidade da coloração da polpa) e uma amostra *backup* de 60g, acondicionadas em frascos de vidro com tampas, recobertos com folha de papel alumínio para evitar ao máximo o contato com a luz para evitar a degradação dos carotenoides. Em seguida as amostras foram congeladas para análise posterior.

Para quantificação do teor de carotenoides foi adotado o procedimento descrito pelo *HarvestPlus Handbook for Carotenoid Analysis* (RODRIGUEZ-AMAYA; KIMURA, 2004). Os pigmentos das raízes trituradas foram extraídos com aproximadamente 50 mL de acetona, por meio da moagem das amostras em homogeneizador Ultra Turrax durante três minutos a fim de fazer com que as células vegetais fossem rompidas e os carotenoides extraídos pela acetona. A massa com a amostra e acetona foi filtrada através de um funil de Buchner com auxílio de uma bomba de vácuo, e o resíduo retido no funil foi lavado com acetona até não apresentar coloração. O extrato contendo apenas o pigmento

e acetona foi reservado no balão de sucção, em seguida transferido para um funil de separação contendo éter de petróleo (a quantidade pode variar a depender da coloração da amostra), e adicionada solução salina lentamente (cerca de 250 mL) para separação das fases. A fase aquosa foi descartada e o procedimento de lavagem com solução salina repetida por cinco vezes até restar apenas éter de petróleo e pigmento. Esse novo extrato foi filtrado em funil com lã de vidro e sulfato de sódio anidro para um balão volumétrico âmbar e completado seu volume com éter de petróleo. Tomou-se uma alíquota de cada amostra para determinação por espectrofotômetro (UV-VIS Thermo Scientific, modelo Genesis 10S) ajustado com uma absorbância a 450 nm.

O TCC foi calculado usando a seguinte fórmula: $TCC(ug.g^{-1}) = \frac{A \times V(ml) \times 10^4}{A_{1cm}^{1\%} \times P(g)}$, em que A é a absorbância; $V(ml)$ é o volume total do extrato em mililitros; $P(g)$ é o peso da amostra em gramas; e $A_{1cm}^{1\%}$ é igual ao coeficiente de extinção do β -caroteno em éter de petróleo (2592).

2.5 Análise dos dados

Inicialmente foi realizada análise de correlação de Pearson entre os dados fenotípicos do TCC e os dados colorimétricos do sistema CIE de cores a fim de verificar a magnitude e o sentido da correlação existente entre estes parâmetros. Foi implementada a análise de componentes principais (PCA) para verificar a diversidade fenotípica dos genótipos com base no TCC e nos índices colorimétricos. A análise PCA também foi usada para verificar a importância dos índices colorimétricos como variáveis preditoras do TCC. O agrupamento dos genótipos foi determinado pelo algoritmo *K-means clustering* e todas estas análises foram realizadas com auxílio do pacote *factoextra* (KASSAMBARA; MUNDT, 2017) implementado no ambiente de programação R versão 4.03 (R CORE TEAM, 2021).

Em seguida os dados colorimétricos foram utilizados na construção de modelos de predição do TCC. Doze modelos de predição derivados de regressão e aprendizado de máquina foram testados: *Linear Regression with Forward Selection* (LRFS), *Linear Regression with Backwards Selection* (LRBS), *Ridge Regression* (RR), *Linear Regression with Stepwise Selection*

(LRSS), *Generalized Linear Model with Stepwise Feature Selection* (GLMSS), *Random Forest* (RF), *Partial Least Squares* (PLS), *The Bayesian Lasso* (BL), *The Bayesian Blasso* (BBL), *Artificial Neural Network* (ANN), *Support vector machine* (SVM), *Classification and regression trees* (CART) (SINGH; BHATIA; SANGWAN, 2007; DOMINGOS, 2012). Os modelos de predição foram implementados no ambiente de programação R versão 4.03 (R CORE TEAM, 2021) com auxílio do pacote *caret* (KUHN, 2008). O esquema de validação cruzada foi realizado em 5 repetições com 6-folds cada. Portanto, 60% das amostras foram usadas como população de treinamento e 40% como população de validação.

O desempenho dos modelos de predição foi avaliado com base na média dos valores do erro quadrático médio (RMSE) e do coeficiente de determinação (R^2), obtidos em cada partição da validação cruzada. O RMSE foi utilizado para analisar os erros da estimativa, sendo a medida da magnitude média dos erros estimados. Possui valor positivo e quanto mais próximo de 0, maior será a qualidade dos valores estimados. É calculado segundo a equação: $RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (E_i - O_i)^2}$, onde: E_i e O_i são os valores estimados e observados, respectivamente, e n é o número de observações. O R^2 varia entre 0 e 1, e quanto mais próximo de 1, melhor será o seu poder de explicação. O R^2 é descrito como a relação que pondera a proporção da variação total da variável dependente que é explicada pela variação da variável independente. É estimado considerando a equação: $R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$, em que: $\sum_{i=1}^n (\hat{y}_i - \bar{y})^2$ corresponde à variação explicada, e $\sum_{i=1}^n (y_i - \bar{y})^2$ corresponde à variação não explicada.

3. RESULTADOS

3.1 Correlação entre teor de carotenoides totais e índices colorimétricos

O TCC apresentou amplitude de $0,30 \mu\text{g g}^{-1}$ a $13,67 \mu\text{g g}^{-1}$ com base no peso fresco das amostras, enquanto a média geral dos 228 genótipos de mandioca foi de $3,56 \mu\text{g g}^{-1}$. Gráficos de dispersão foram construídos para

avaliar a distribuição dos dados e estimar a correlação de Pearson entre o TCC e os índices colorimétricos extraídos das imagens digitais (Figura 1). Alta correlação positiva foi identificada entre TCC e as variáveis b^* ($r = 0,90$) e *chroma* ($r = 0,89$), enquanto que correlações medianas e negativas foram identificadas entre TCC e os índices a^* ($r = -0,53$), *Hue* ($r = -0,59$) e Luminosidade ($r = -0,66$). Por outro lado, o parâmetro L^* não apresentou correlação significativa com TCC.

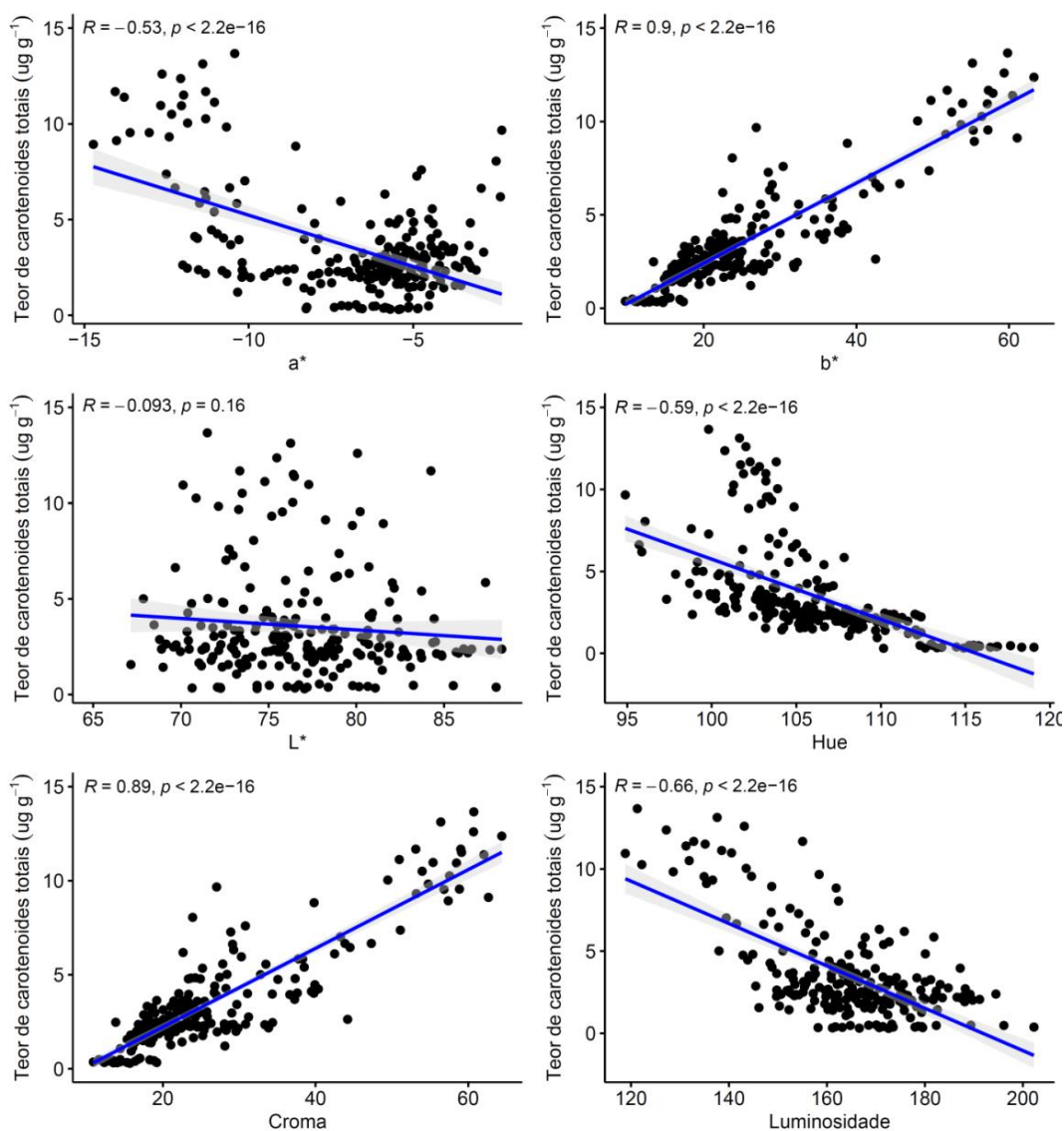


Figura 1. Gráficos de dispersão entre o teor de carotenoides totais e índices colorimétricos. Os valores numéricos representam os coeficientes de correlação de Pearson entre os índices colorimétricos e o teor de carotenoides totais. A linha em azul representa a isolinha 1:1.

3.2 Diversidade fenotípica e agrupamento dos genótipos de mandioca com base no teor de carotenoides totais e nos índices colorimétricos

A estruturação populacional com base nos dados fenotípicos de TCC e índices colorimétricos foi avaliada com base na análise de componentes principais (PCA). De acordo com o algoritmo *K-means clustering* foram identificados cinco grupos com base na informação da soma de quadrados da variação dentro de grupos (Figura 2).

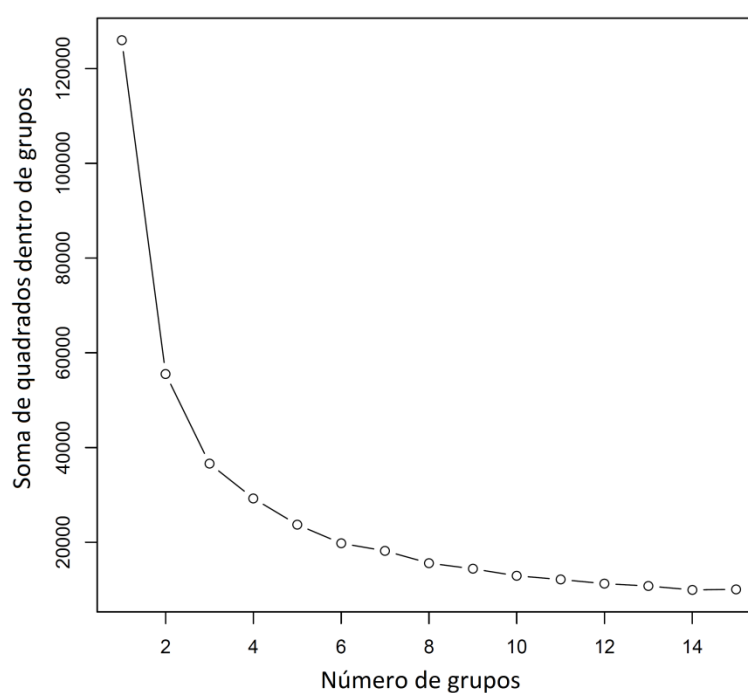


Figura 2. Soma de quadrados da variância dentro de grupos em 228 genótipos de mandioca biofortificados com base em dados fenotípicos do teor de carotenoides totais ($\mu\text{g.g}^{-1}$) e índices colorimétricos (CIELAB). O critério de agrupamento foi realizado com base no algoritmo *K-means clustering*.

Os dois primeiros componentes principais explicaram 89% da variância total do conjunto de dados da população amostral e, portanto, permitem uma boa representação do agrupamento dos genótipos de mandioca com base nos parâmetros avaliados (Figura 3). Houve ampla variação fenotípica entre os genótipos de mandioca para TCC e também para os índices colorimétricos. O TCC e os índices colorimétricos b^* e $croma$ foram as variáveis que mais contribuíram para dispersão dos genótipos. De forma semelhante ao que foi

observado nas análises de correlações, os índices b^* e *croma* foram os que apresentaram maior associação com TCC, e por isso possuem alta importância relativa para a predição deste fenótipo. Por outro lado, o índice L^* foi o que menos contribuiu para variação fenotípica dos genótipos de mandioca.

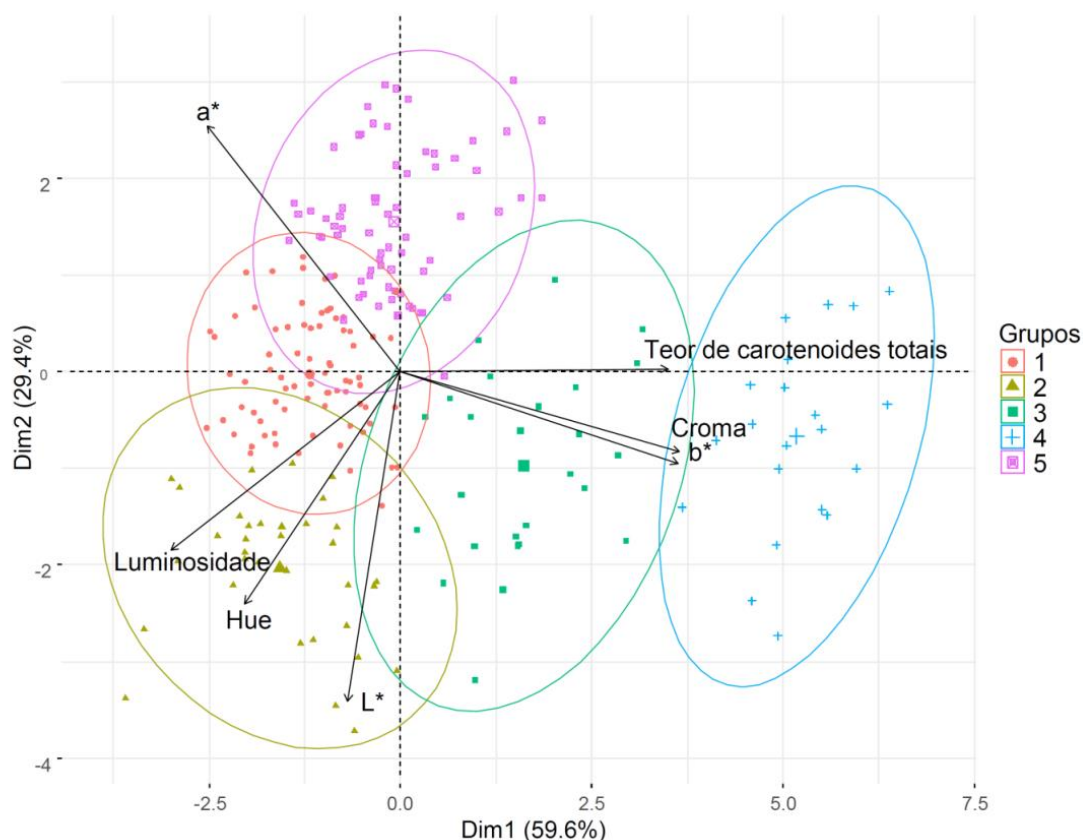


Figura 3. Análise biplot de componentes principais (PCA) com base em dados fenotípicos do teor de carotenoides totais ($\mu\text{g.g}^{-1}$) e índices colorimétricos (CIELAB) avaliados em 228 genótipos de mandioca biofortificados.

Com base na PCA os genótipos de mandioca foram agrupados em cinco grupos. Os grupos 1, 2, 3, 4 e 5 reuniram 76, 34, 26, 21 e 71 genótipos, respectivamente. Os grupos 1 e 2 foram constituídos por genótipos com baixos valores de TCC ($1,91 \mu\text{g.g}^{-1}$ e $1,82 \mu\text{g.g}^{-1}$, respectivamente), b^* e *croma*, além dos maiores valores de L^* , luminosidade e *Hue* (Figura 4). Os grupos 3 e 5 foram constituídos por genótipos de mandioca com polpa amarela, mas com teores medianos de TCC ($4,82 \mu\text{g.g}^{-1}$ e $3,60 \mu\text{g.g}^{-1}$, respectivamente) e por valores medianos de luminosidade (variação de 158 a 159). Por outro lado, estes dois grupos diferem quanto aos valores a^* ($-10,23$ e $-4,60$ nos grupos 3 e 5, respectivamente), b^* ($36,48$ e $22,20$ nos grupos 3 e 5, respectivamente),

L^* (78,60 e 73,10 nos grupos 3 e 5, respectivamente), Hue (105,59 e 101,87 nos grupos 3 e 5, respectivamente) e $croma$ (37,93 e 22,72 nos grupos 3 e 5, respectivamente).

O grupo 4 foi composto pelos genótipos de mandioca com maior TCC (média de $10,74 \mu\text{g}\cdot\text{g}^{-1}$). Em relação aos índices colorimétricos, o grupo 4 foi caracterizado por apresentar alto $croma$ (média de 56,96) e b^* (média de 55,57), baixos valores de Hue (média de 102,59), a^* (média de -12,38) e luminosidade (média de 136,04) e valores medianos de L^* (média de 76,14).

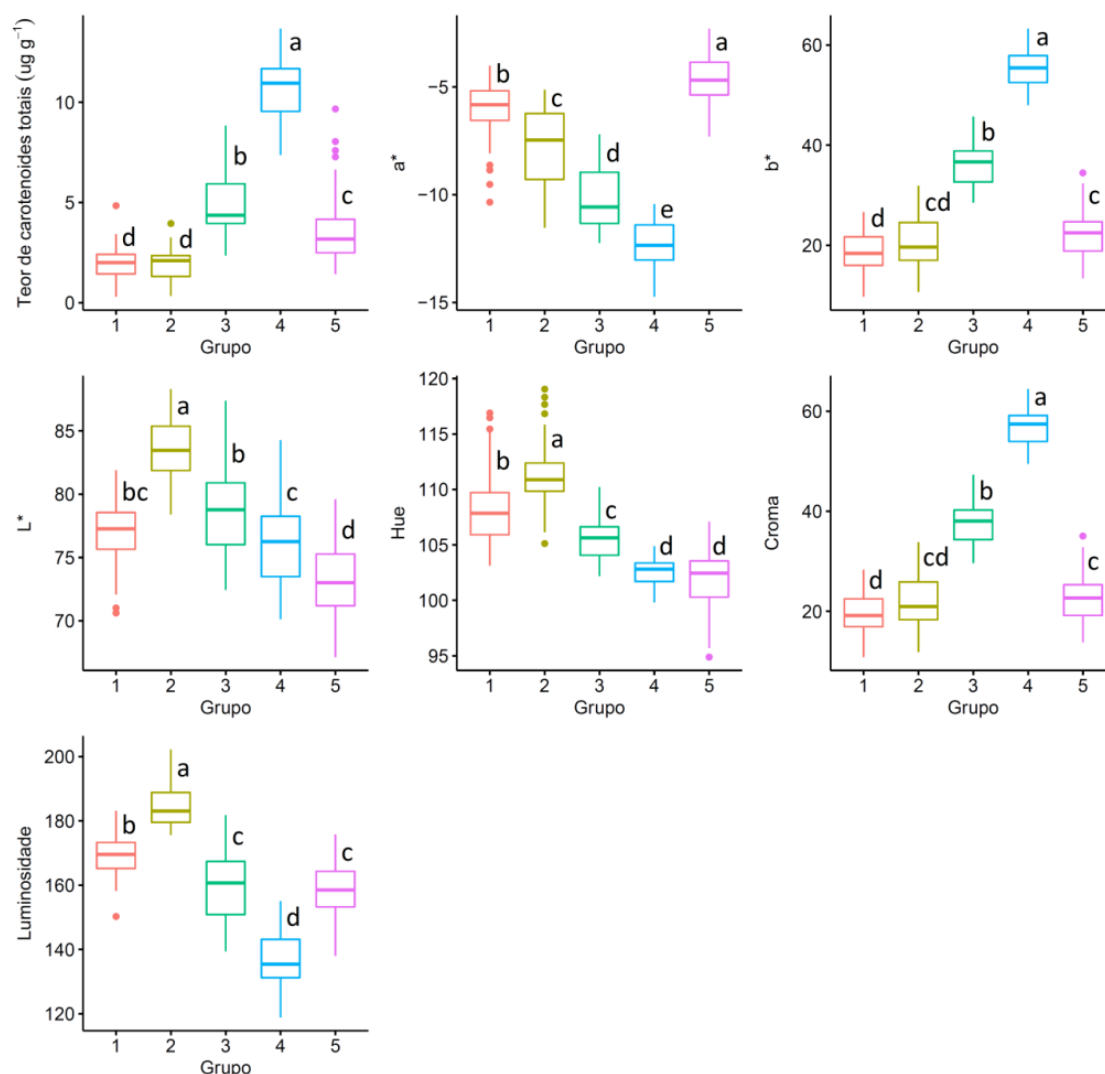


Figura 4. Boxplot dos valores médios do teor de carotenoides totais e dos índices colorimétricos, para cada agrupamento dos 228 genótipos de mandioca formado com base na análise de PCA. Letras diferentes representam diferenças significativas entre grupos de acessos, $p < 0.05$ pelo teste Tukey HSD.

3.3 Predição do teor de carotenoides totais utilizando índices colorimétricos

O desempenho dos modelos de predição para TCC em raízes de mandioca utilizando dados colorimétricos obtidos de imagens digitais foi avaliado considerando o modelo completo (todas as variáveis para predição de TCC) e o modelo reduzido, na qual apenas as variáveis com importância relativa maior ou igual a 50% para a predição de TCC foram mantidas nos diferentes modelos avaliados.

Os índices colorimétricos b^* e *croma* com importância relativa de 100% foram as variáveis que mais contribuíram para a predição de TCC na maioria dos modelos (Figura 5). A única exceção ocorreu no modelo PLS em que os índices b^* e a^* apresentaram importância relativa de 100%. Os índices Luminosidade e *Hue* apresentaram importância entre 50 e 70% dependendo do modelo analisado. Para todos os modelos o índice L^* apresentou importância inferior a 50% e por isso foi excluído nos modelos reduzidos. Portanto, a maioria dos modelos reduzidos incluiu as variáveis a^* , b^* , *croma*, *Hue* e Luminosidade, à exceção do RF cujo modelo reduzido foi representado apenas por b^* , *Hue* e *croma*. Apesar do modelo reduzido possuir apenas as variáveis como maior capacidade preditiva do TCC, não houveram diferenças na capacidade preditiva em comparação com o modelo completo (Tabela 1). Apenas os modelos RF e PLS apresentaram uma pequena diferença na acurácia de predição, ou seja, $R^2=0,91$ e $0,90$ (RF) e $R^2=0,94$ e $0,93$ (PLS) nos modelos completos e reduzidos, respectivamente.

De modo geral a acurácia dos modelos de predição (completo e reduzido) foi elevada, com R^2 variando de $0,90$ (modelo SVM) a $0,94$ (modelo ANN) (Tabela 1). Apenas o modelo de predição CART apresentou acurácia de predição do TCC abaixo de $0,90$. Por outro lado, o melhor modelo de predição foi o ANN, por apresentar a melhor capacidade preditiva ($R^2=0,94$), associada ao menor erro nas estimativas do TCC (RMSE= $0,24$). Os modelos baseados em regressão linear (LRFS, LRBS, LRSS, GLMSS) bem como os modelos PLS, BL, BBL e RR apresentaram desempenho semelhante (RMSE= $0,23$; $R^2=0,93$). Apesar dos modelos RF e SVM apresentarem alta capacidade

preditiva ($R^2=0,91$ e $0,90$, respectivamente), eles também apresentaram maiores valores de RMSE ($0,32$ e $0,33$, respectivamente).

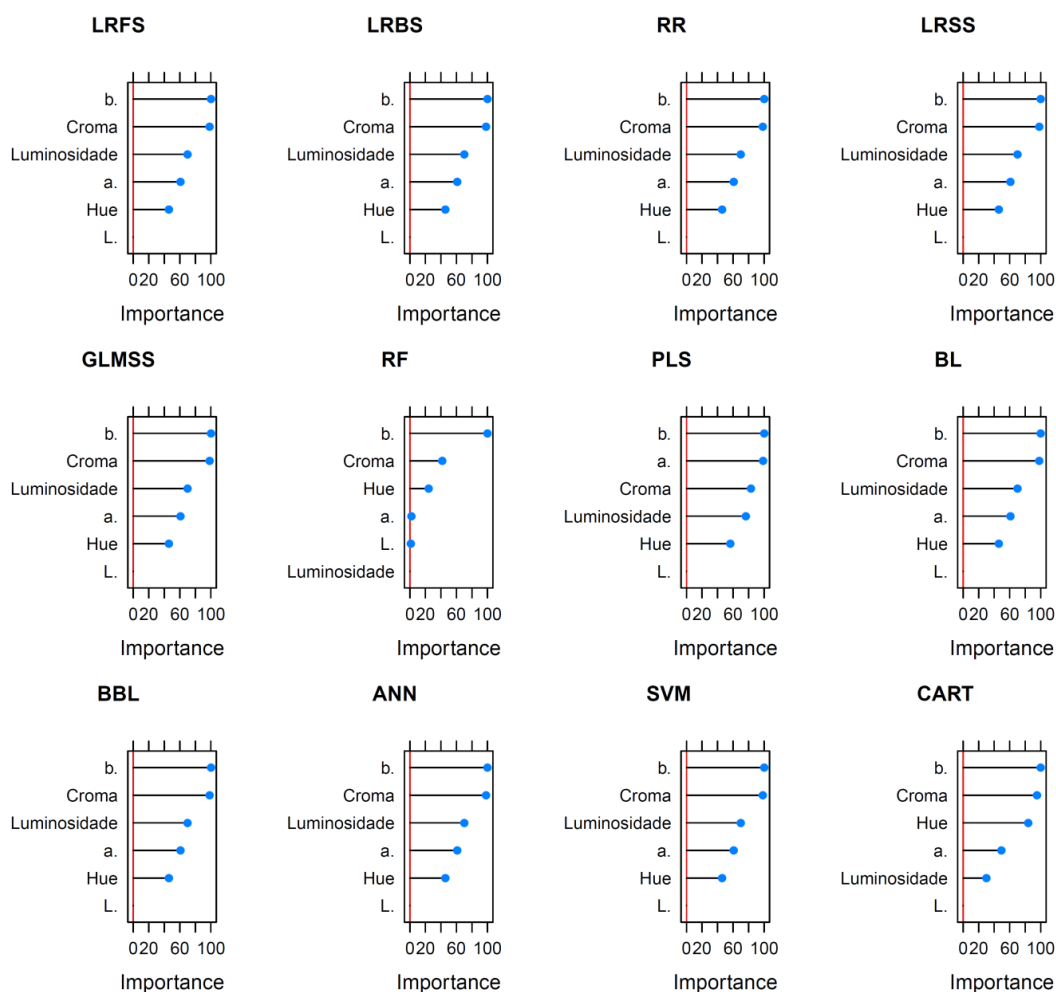


Figura 5. Importância relativa de índices colorimétricos para predição do teor de carotenoides totais usando 12 modelos de predição. *Linear Regression with Forward Selection* (LRFS), *Linear Regression with Backwards Selection* (LRBS), *Ridge Regression* (RR), *Linear Regression with Stepwise Selection* (LRSS), *Generalized Linear Model with Stepwise Feature Selection* (GLMSS), *Random Forest* (RF), *Partial Least Squares* (PLS), *The Bayesian Lasso* (BL), *The Bayesian Blasso* (BBL), *Artificial Neural Network* (ANN), *Support vector machine* (SVM), *Classification and regression trees* (CART).

O gráfico de dispersão dos valores observados e preditos para TCC mostra claramente o maior ajuste do modelo ANN ($R^2=0,94$) seguido pelos modelos LRFS, LRBS, LRSS, GLMSS, PLS, BL, BBL e RR ($R^2= 0,93$) independentemente do *fold* da validação (Figura 6). Por outro lado, o modelo CART ($R^2=0,82$), RF e SVM ($R^2=0,90$) apresentaram maior dispersão dos

dados preditos e, portanto, menor ajuste para predição de TCC em raízes de mandioca.

De modo geral, para a maioria dos modelos, as predições de TCC com valores entre 0,50 e 5,0 $\mu\text{g g}^{-1}$ foi bastante acurada. Entretanto, acima deste valor, os gráficos de regressão mostram uma maior dispersão dos dados. Isso possivelmente ocorreu pelo fato de que a maioria dos genótipos (82%) apresentaram TCC abaixo de 5,0 $\mu\text{g g}^{-1}$, representando assim a maior parte das amostras.

Tabela 1. Desempenho de diferentes modelos de predição para teor de carotenoides totais em raízes de mandioca utilizando índices colorimétricos obtidos de imagens digitais e considerando o modelo completo (todas as variáveis) e modelo reduzido (apenas variáveis com mais de 50% de importância relativa).

Modelo	Completo		Reduzido	
	RMSE ^a	R ^{2a}	RMSE	R ²
Linear Regression with Forward Selection	0,26	0,93	0,26	0,93
Linear Regression with Backwards Stepwise Selection	0,26	0,93	0,26	0,93
Ridge Regression (method = 'ridge')	0,26	0,93	0,26	0,93
Linear Regression with Stepwise Selection	0,26	0,93	0,26	0,93
Generalized Linear Model with Stepwise Feature Selection	0,26	0,93	0,26	0,93
Random Forest (method = 'Rborist')	0,31	0,91	0,32	0,90
Partial Least Squares	0,26	0,94	0,26	0,93
The Bayesian lasso (method = 'lasso')	0,26	0,93	0,26	0,93
The Bayesian blasso (method = 'blasso')	0,26	0,93	0,26	0,93
Artificial Neural Network	0,24	0,94	0,24	0,94
Support vector machine	0,33	0,90	0,33	0,90
Classification and regression trees (Decision Trees)	0,43	0,81	0,43	0,81

^aR²: coeficiente de determinação e RMSE: raiz do erro quadrático médio.

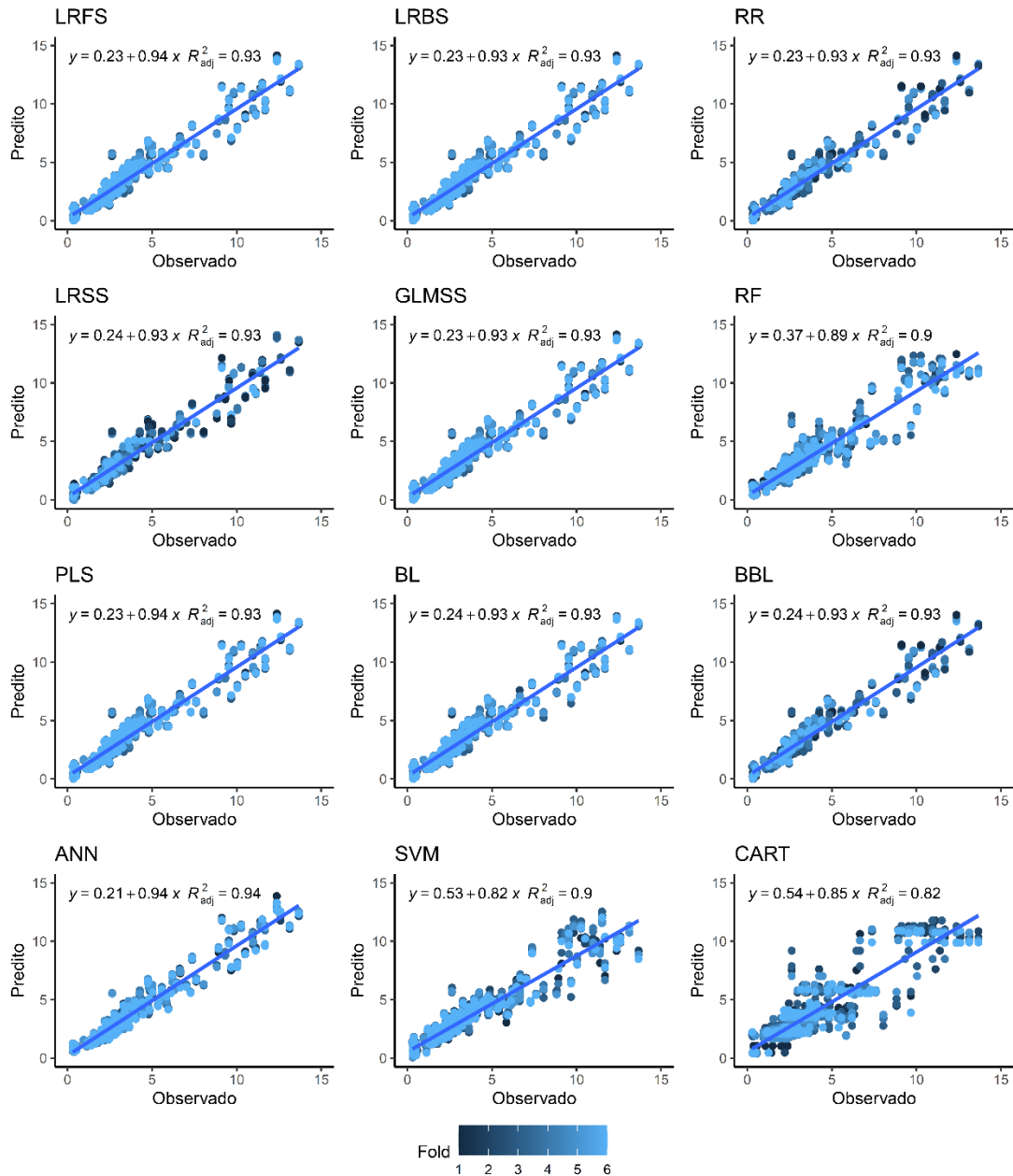


Figura 6. Relação entre os valores observados e preditos para teor de carotenoides totais de raízes de mandioca. A predição foi realizada com base em 12 diferentes modelos com base em validação cruzada. *Linear Regression with Forward Selection* (LRFS), *Linear Regression with Backwards Selection* (LRBS), *Ridge Regression* (RR), *Linear Regression with Stepwise Selection* (LRSS), *Generalized Linear Model with Stepwise Feature Selection* (GLMSS), *Random Forest* (RF), *Partial Least Squares* (PLS), *The Bayesian Lasso* (BL), *The Bayesian Blasso* (BBL), *Artificial Neural Network* (ANN), *Support vector machine* (SVM), *Classification and regression trees* (CART). Cada fold da validação cruzada está representado em diferentes cores. Os dados numéricos incluídos nos gráficos representam: equações lineares (y) e coeficiente de determinação (R^2). A linha representa a isolinha 1:1.

4. DISCUSSÃO

4.1. Variação e correlação entre o teor de carotenoides totais e índices colorimétricos

O TCC nos 228 genótipos de mandioca analisados apresentou amplitude de $0,30 \mu\text{g g}^{-1}$ a $13,67 \mu\text{g g}^{-1}$. Esta variação no TCC é coerente com outros relatos em mandioca, a exemplo da triagem de 1789 genótipos de mandioca do Centro Internacional de Agricultura Tropical (CIAT, Colômbia), cuja variação no TCC foi de 1,02 a $10,40 \mu\text{g g}^{-1}$ (CHÁVEZ et al., 2005). Entretanto, em populações de melhoramento especificamente voltadas para o aumento do TCC, esta variação pode ser ainda mais ampla a exemplo da variação de $0,2 \mu\text{g g}^{-1}$ a $25,5 \mu\text{g g}^{-1}$ observada por Sánchez et al. (2014).

A variação observada no TCC constitui-se em uma ampla faixa de variação para construção de modelos de predição com alta acurácia para esta característica. E, portanto, neste estudo foi analisado o potencial de uso da fenotipagem rápida com base imagens digitais e a capacidade de estimar o TCC nas raízes de mandioca à partir de dados colorimétricos utilizando o sistema de cores CIELAB (ITLE; KABELKA, 2009; AFONSO et al., 2017).

A presença dos pigmentos carotenoides contribui para intensidade da coloração da polpa das raízes de mandioca (CEBALLOS et al., 2013; CEBALLOS et al., 2017), conferindo uma correlação positiva entre estas duas características (AFONSO et al., 2017). Assim, primeiramente foi avaliada a relação entre TCC e as variáveis relacionadas à cor, sendo observada alta correlação entre TCC e b^* ($r=0,90$). Esta alta correlação era esperada uma vez que os valores positivos do eixo b^* correspondem a cor amarela. Este resultado está de acordo com estudos em outras espécies onde foi verificado que o acúmulo de carotenoides indica uma tendência de aumento de b^* , como verificado em diversas variedades de tangerina Kinnow (*Citrus reticulata* Blanco) (HADIMANI; MITTAL, 2019) e em amostras de milho extrusadas (CUETO et al., 2017). Em mandioca, Afonso et al. (2017) também constataram que os maiores valores de b^* estavam associados às amostras de mandioca com maior conteúdo de carotenoide. Esse comportamento se deve à correlação entre o parâmetro b^* com o β -caroteno (KLJAK; GRBEŠA;

KAROLYI, 2014), uma vez que o β -caroteno é o carotenoide mais abundante na mandioca (CEBALLOS et al., 2017).

O mesmo padrão de correlação foi observado entre TCC e *croma* ($R=0,89$) (Figura 1). O parâmetro *croma* indica a intensidade da coloração (PATHARE; OPARA; AL-SAID, 2013), e portanto, pode-se afirmar que quanto maior o teor de carotenoides totais maior será a intensidade da coloração da polpa de mandioca. Itle; Kabelka (2009) encontraram correlações semelhantes entre b^* e *croma* ao analisarem o teor de carotenoides individuais e TCC em abóboras (*Cucurbita* spp). Além disso, estes autores observaram que o ângulo *Hue* apresentou correlação negativa com TCC ($r=-0,83$), de forma semelhante ao observado no presente estudo em mandioca ($r=-0,59$), sugerindo que à medida que os ângulos diminuem, as concentrações de carotenoides aumentam. Além do mais, Itle; Kabelka (2009) identificaram fortes correlações entre os índices colorimétricos e o conteúdo de carotenoides, e afirmaram que essas correlações podem ser úteis na seleção indireta para alto teor de carotenoides em populações melhoradas de abóbora, sendo de fácil implementação e ao mesmo tempo de baixo custo.

4.2 Agrupamento da diversidade fenotípica do teor de carotenoides totais com base nos índices colorimétricos

Os dois componentes principais explicaram quase 90% da variância fenotípica, e com isso foram capazes de demonstrar a dispersão dos genótipos de mandioca e a formação de cinco grupos com características similares em termos de TCC e índices colorimétricos. Especificamente os grupos 3 e 4 reuniram os acessos com polpa amarela e conseqüente maiores teores de TCC, b^* e *croma*, visto que esses dois índices colorimétricos estão positivamente correlacionados com TCC

O potencial de agrupamento de genótipos com base nos índices colorimétricos tem sido recentemente analisado em diversas espécies e até mesmo em mandioca. Afonso et al. (2017) avaliaram o potencial de agrupamento de diferentes amostras de raízes de mandioca com base nos índices do espaço de cores CIELAB, obtidos por meio de um colorímetro. Os autores identificaram que as amostras com maior TCC apresentaram maiores

valores de b^* . Os autores também reportaram agrupamento dos genótipos via PCA com base na coloração da polpa da raiz e no TCC, utilizando dados colorimétricos.

Moresco et al. (2015), utilizaram PCA e outros métodos de agrupamentos a partir da matriz de dados espectrofotométrico aplicado na região UV-Vis (400-500nm) em amostras de raízes de mandioca com cor de polpa creme, amarela e avermelhada. Nesta estratégia, os dois primeiros componentes representaram 99,97% da variância, revelando claramente três grupos de acordo com os teores de carotenoides. Estes autores atribuem o agrupamento dos genótipos ao teor de carotenoides e a discrepância do genótipo de polpa avermelhada à presença de licopeno em quantidades relevantes, detectado pela análise cromatográfica.

A eficiência do PCA em distinguir diferentes amostras também foi descrita por Nankar et al. (2020), cujo estudo reportou a diversidade morfométrica e colorimétrica dos frutos da coleção de pimenta (*Capsicum* spp.) dos Balcãs e o agrupamento via PCA. Uma das características avaliadas que mais contribuiu para a variabilidade das pimentas foi a cor do fruto, estimadas pelos parâmetros do sistema CIELAB, obtidos pelo software Tomato Analyzer. Zaouay et al. (2012) reportaram a similaridade entre as variedades de romã (*Punica granatum* L.) quanto à atividade antioxidante e as propriedades físico-químicas de frutos. Esses autores verificaram a contribuição significativa no agrupamento realizado pela análise PCA do índice a^* mensurado no suco da fruta e sua correlação com outros atributos importantes para a cultura, a exemplo do teor de fenólicos totais e conteúdo total de antocianinas.

4.2. Desempenho dos modelos de predição do TCC

Os diferentes modelos de predição baseados em aprendizagem de máquina apresentaram elevada capacidade de predição do TCC em raízes de mandioca, em função dos altos valores de R^2 e baixos valores de RMSE. Entretanto, o modelo baseado em redes neurais artificiais (ANN) foi o modelo mais acurado em comparação aos demais por apresentar melhor ajuste do RMSE (0,24) e R^2 (0,94). Resultados similares também foram observados por Sanaeifar; Jafari; Golmakani (2018), que identificaram alta capacidade preditiva

do teor de carotenoides em azeite de oliva (*Olea europaea*) utilizando ANN ($R^2=0,94$) e SVM ($R^2=0,96$). Em tangerina (*Citrus reticulata Blanco*), os modelos PLS ($R^2=0,96$) e regressão linear de terceira ordem ($R^2=0,95$) foram capazes de prever a maturação (*citrus color index*) com alta acurácia (HADIMAN; MITTAL, 2019). González et al. (2020), apresentaram uma metodologia para predição do teor de carotenoides individuais a partir de parâmetros de análise de imagens digitais de amostras de pólen de espécies das famílias Brassicaceae, Myrtaceae, Fabaceae, com base em regressão linear múltipla (MLR), com altos R^2 (variação de 0,76 a 0,89).

ANN é um modelo que utiliza técnicas inspiradas no cérebro e na maneira como ele aprende e processa informações (WANG; SUMMERS, 2012). O modelo ANN, assim como os demais utilizam aprendizagem supervisionada, que visa fornecer a previsão de uma variável de saída de acordo com as variáveis de entrada conhecidas (LIAKOS et al., 2018). Os modelos de aprendizagem de máquina tem sido uma ferramenta eficiente bastante utilizada em diferentes abordagens na agricultura, para desvendar, quantificar e compreender processos intensivos de dados (LIAKOS et al., 2018). Entretanto, a capacidade preditiva depende do conjunto de dados em análise e por esse motivo é interessante avaliar diferentes abordagens a fim de explorar diferenças nos algoritmos que podem beneficiar as predições (CARMO et al., 2020).

A busca por metodologias de fenotipagem rápida tem sido prioridade em programas de melhoramento de mandioca. Portanto, vários grupos têm buscado alternativas para encurtar os ciclos de seleção, aumentar o número de genótipos avaliados e diminuir custos associados com a fenotipagem de características que são difíceis de mensuração a exemplo do teor de carotenoides. Mesmo com a implementação de metodologias de fenotipagem rápida, poucos estudos têm se dedicado à validação dos modelos de predição de características associadas à qualidade nutricional como o teor de matéria seca e carotenoides em mandioca com base em análise de imagens. Os modelos com boa acurácia de predição podem permitir a seleção precoce apenas dos genótipos mais interessantes, permitindo a otimização da seleção.

O uso de variáveis preditivas de fácil mensuração, como os dados colorimétricos extraídos das imagens digitais, é um ponto chave para acelerar a triagem dos genótipos em programas de melhoramento que visam o *screening* de milhares de indivíduos anualmente. Esta abordagem demonstrou que as variáveis b^* e *croma* foram fortemente correlacionadas com o TCC em raízes de mandioca (Figura 1). Este resultado permite simplificar ainda mais o protocolo de fenotipagem para o TCC em raízes de mandioca utilizando as imagens digitais, pois permite a redução de variáveis, sem perder a capacidade preditiva da característica.

Em mandioca, outros autores também utilizaram técnicas colorimétricas para predição do TCC em raízes. Afonso et al. (2017) demonstraram que a medição da cor a partir dos dados colorimétricos usando o sistema de cores CIELAB, pode ser usada como um método rápido e não destrutivo para calibrar o TCC de raízes com erro de predição aceitável. Estes autores utilizaram dados CIELAB como entrada para os vários modelos de aprendizagem de máquina e dados de TCC determinado por espectrofotometria e encontraram valores de $R^2 \sim 0.60$. Moresco et al. (2017), também utilizaram a técnica colorimétrica associada a UV-vis / HPLC para predição do TCC em raízes de mandioca. Estes últimos autores encontraram valores de $R^2 > 0,40$ e $RMSE < 9,99$ nos modelos PLS, SVM e *Elastic Net*. Portanto, as abordagens utilizadas no presente estudo permitiram a obtenção de maior capacidade preditiva e menor erro em comparação com relatos semelhantes na literatura. Uma possível explicação para isso seria o fato de que Afonso et al. (2017) e Moresco et al. (2017) utilizaram colorímetro na obtenção dos índices colorimétricos CIELAB. Especula-se que as imagens digitais permitem maior precisão na análise de cores, uma vez que capturam todos os pixels das imagens das raízes, enquanto que os colorímetros medem apenas uma porção específica dessa amostra.

O software utilizado no presente estudo para análise das imagens RGB (*Tomato Analyzer*) também apresentou alto potencial para quantificação de características de difícil mensuração por métodos convencionais de fenotipagem, a exemplo da espessura da parede do fruto e espessura do pericarpo em frutos de pimenta (*Capsicum spp.*) avaliados no germoplasma

coletado em diferentes regiões dos Balcãs (NANKAR et al., 2020). O *Tomato Analyzer* permite estimar um grande número de características associadas ao formato e cores derivadas de imagens digitalizadas de seções de frutas (FIGÀS et al., 2015) ou mesmo de raízes como foi o caso do estudo atual.

De uma forma geral, o desenvolvimento de modelos de predição baseados em diferentes abordagens tem demonstrado sua grande utilidade e confiabilidade para estimar o TCC em mandioca. Um dos primeiros estudos para predição do teor de carotenoides em mandioca foi desenvolvido com base em dados NIRS (SÁNCHEZ et al., 2014), cuja capacidade preditiva foi de 0,92 e 0,93 para TCC e β -caroteno, respectivamente, utilizando regressão PLS. Em outro estudo, também utilizando dados NIRS, Belalcazar et al. (2016) demonstraram pela primeira vez o uso prático das predições do teor de carotenoides para seleção nos programa de melhoramento da espécie. Com base no uso de regressão LOCAL, estes autores relataram $R^2=0,74$ para o teor de β -caroteno. Mais recentemente Ikeogu et al. (2017, 2019), também utilizaram dados NIRS, mas com diferentes modelos de predição, a exemplo do *modified partial least squares* (MPLS) e RF. Estes últimos autores relataram para elevada capacidade preditiva do teor de β -caroteno ($R^2=0,99$).

No presente estudo, a fenotipagem via imagens digitais também demonstrou sua elevada capacidade preditiva ($R^2=0,94$ e $RMSE=0,24$) para predição do TCC com base no modelo ANN, e, portanto, esta é uma nova abordagem para mensuração desta característica em mandioca, com custo muito menor na aquisição e manutenção de equipamentos. A implementação de sistema de visão computacional baseado em imagens RGB para detectar características colorimétricas em tomate, utilizando equipamentos de baixo custo, apresentou excelente capacidade de classificação (WAN et al., 2018). Ye; Izawa; Zhang (2018), também desenvolveram um aplicativo para determinação rápida do teor de licopeno nos frutos de tomate e classificação dos diferentes estágios de maturação com base nas cores, a partir dos parâmetros L^* a^* b^* , obtidos de imagens RGB. Portanto, a utilização de técnicas de análise de imagens digitais, com base em softwares livres e câmeras de baixo custo, permitem obter modelos de predição acurados e confiáveis para quantificação do TCC em mandioca. Todos esses aspectos

otimizam o tempo de análise, recursos humanos e financeiros, sem perder a confiabilidade nos resultados.

4.3. Perspectivas futuras

O potencial de implementação da fenotipagem rápida baseada em imagens digitais para predição do TCC foi demonstrada neste estudo. Além disso, a análise de imagens permite analisar e incorporar outras características de importância agrônômica na cultura da mandioca a exemplo da avaliação de características relacionadas à estrutura do dossel e rendimento da raiz, avaliação da ligação entre arquitetura da raiz e concentração de micronutrientes (zinco e cálcio) e estimativa do crescimento e desempenho nutricional da mandioca sob irrigação deficitária e fertirrigação com potássio (SELVARAJ et al., 2020; BUSENER et al., 2020; WASONGA et al., 2021). Portanto, estas variáveis podem vir a ser utilizadas na composição de índices de seleção multi-características para maior otimização dos programas de melhoramento de mandioca. Além disso, o uso conjunto desta abordagem com ferramentas genômicas de seleção pode acelerar o desenvolvimento de variedades de mandioca com maior teor de carotenoides totais.

Essa foi uma primeira etapa para a construção de um banco de dados de imagens digitais para predição do TCC em mandioca. Os próximos passos envolverão a inclusão de amostras de raízes com maior TCC para que seja possível reduzir o erro de predição, sobretudo em amostras com mais de $5,0 \mu\text{g g}^{-1}$. Além disso, amostras de raízes cultivadas em diferentes ambientes de cultivo serão avaliadas para inclusão do efeito da interação genótipo \times ambiente nos modelos de predição para que eles possam ter uso mais generalizado nos diferentes programas de melhoramento genético. Todos estes aperfeiçoamentos tem como foco principal o aumento da capacidade preditiva do TCC e a implementação prática da fenotipagem de alto rendimento para *screening* de centenas de amostras diárias. A expectativa é que as metodologias de fenotipagem baseada em imagens permitam melhorar a qualidade e precisão dos dados coletados em diversas avaliações fenotípicas, contribuindo assim para aumento no diferencial de seleção e no coeficiente de herdabilidade, com efeito direto no ganho genético (CORTES et al., 2017).

5. CONCLUSÃO

O uso de dados colorimétricos do espaço CIELAB obtidos a partir da análise de imagens digitais RGB é uma alternativa econômica, rápida e eficaz, que demonstrou excelente capacidade preditiva para o teor de carotenoides totais em raízes de mandioca. Os índices b^* e *croma* do modelo CIE de cores estão fortemente correlacionados com TCC e podem ser usados com precisão para avaliação desta característica em mandioca. Também foram as variáveis que mais influenciaram na acurácia dos modelos preditivos, e que mais contribuiu para o agrupamento dos genótipos com níveis mais altos de TCC baseado na PCA.

O modelo ANN apresentou a maior capacidade preditiva do TCC em raízes de mandioca em comparação com os demais algoritmos, chegando a apresentar acurácia e erros de predição comparável com abordagens mais sofisticadas a exemplo do NIRS. Dessa forma, esta primeira prova de conceito demonstrou o elevado potencial do uso de imagens digitais para implementação prática da fenotipagem de alto rendimento em mandioca.

6. REFERÊNCIAS BIBLIOGRÁFICAS

AFONSO, T.; MORESCO, R.; UARROTA, V. G.; NAVARRO, B. B.; NUNES, E. D. C.; MARASCHIN, M.; ROCHA, M. UV-Vis and CIELAB based chemometric characterization of *Manihot esculenta* carotenoid contents. **Journal of integrative bioinformatics**, v. 14, n. 4, p. 20170056, 2017. DOI:10.1515/jib-2017-0056.

ALAMU, E. O.; MAZIYA-DIXON, B.; FELDE, T. Z.; KULAKOW, P.; PARKES, E. Application of near infrared reflectance spectroscopy in screening of fresh cassava (*Manihot esculenta* crantz) storage roots for provitamin A carotenoids. Proceedings of the 18th International Conference on Near Infrared Spectroscopy. Chichester, **IM Publications**, p. 91–97, 2019. <https://doi.org/10.1255/nir2017.091>.

BAI, G.; JENKINS, S.; YUAN, W.; GRAEF, G. L.; GE, Y. Field-based scoring of soybean iron deficiency chlorosis using RGB imaging and statistical learning. **Frontiers in Plant Science**, v. 9, p. 1002, 2018. DOI: 10.3389/fpls.2018.01002.

BELALCAZAR, J.; DUFOUR, D.; ANDERSSON, M. S.; PIZARRO, M.; LUNA, J.; LONDOÑO, L.; MORANTE, N.; JARAMILLO, A. M.; PINO, L.; LÓPEZ-LAVALLE, B. L. A.; DAVRIEUX, F.; TALSMA, E. F.; CEBALLOS, H. High-throughput phenotyping and improvements in breeding cassava for increased carotenoids in the roots. **Crop Science**, v. 56, n. 6, p. 2916-2925, 2016. <https://doi.org/10.2135/cropsci2015.11.0701>.

BURNS, A.; GLEADOW, R.; CLIFF, J.; ZACARIAS, A.; CAVAGNARO, T. Cassava: the drought, war and famine crop in a changing world. **Sustainability**, v. 2, n. 11, p. 3572-3607, 2010. <https://doi.org/10.3390/su2113572>

BUSENER, N.; KENGGANNA, J.; SAENGWILAI, P. J.; BUCKSCH, A. Image-based root phenotyping links root architecture to micronutrient concentration in cassava. **Plants People Planet**, v. 2, n. 6, p. 678-687, 2020. <https://doi.org/10.1002/ppp3.10130>.

CARMO, C. D. do; SOUSA, M. B.; PEREIRA, J. D. S.; CEBALLOS, H.; de OLIVEIRA, E. J. Identification of waxy cassava genotypes using fourier-transform near-infrared spectroscopy. **Crop Science**, v. 60, n. 2, p. 883-895, 2020. <https://doi.org/10.1002/csc2.20102>.

CEBALLOS, H.; DAVRIEUX, F.; TALSMA, E. F.; BELALCAZAR, J.; CHAVARRIAGA, P.; ANDERSSON, M. S. Carotenoids in cassava roots. In: **Carotenoids**. InTech, p. 189-221, 2017. <http://dx.doi.org/10.5772/intechopen.68279>.

CEBALLOS, H.; HERSHEY, C.; BECERRA-LÓPEZ-LAVALLE, L. A. New approaches to cassava breeding. **Plant Breeding Reviews**, v. 36, p. 427–504, 2012. DOI:10.1002/9781118358566

CEBALLOS, H.; MORANTE, N.; SÁNCHEZ, T.; ORTIZ, D.; ARAGÓN, I.; CHÁVEZ, A. L.; PIZARRO, M.; CALLE, F.; DUFOUR, D. Rapid cycling recurrent selection for increased carotenoids content in cassava roots. **Crop Science**, v. 53, n. 6, p. 2342-2351, 2013. <https://doi.org/10.2135/cropsci2013.02.0123>

CHÁVEZ, A. L.; SÁNCHEZ, T.; JARAMILLO, G.; BEDOYA, J. M. I.; ECHEVERRY, J.; BOLAÑOS, E.A.; IGLESIAS, C. A. Variation of quality traits in cassava roots evaluated in landraces and improved clones. **Euphytica**, v. 143, p. 125–133, 2005. <https://doi.org/10.1007/s10681-005-3057-2>

COMMISSION INTERNATIONALE DE L'ÉCLAIRAGE (CIE). Recommendations on uniform color spaces, color-difference equations, psychometric color terms. **Paris: CIE**, 1978.

CORTES, D. F. M.; CATARINA, R. S.; BARROS, G. B. D. A.; ARÊDES, F. A. S.; SILVEIRA, S. F. D.; FERREGUETTI, G. A.; RAMOS, H. C. C.; VIANA, A. P.; PEREIRA, M. G. Model-assisted phenotyping by digital images in papaya breeding program. **Scientia Agricola**, v. 74, n. 4, p. 294-302, 2017. <https://doi.org/10.1590/1678-992X-2016-0134>.

CUETO, M.; FARRONI, A.; SCHOENLECHNER, R.; SCHLEINING, G.; BUERA, P. Carotenoid and color changes in traditionally flaked and extruded products. **Food chemistry**, v. 229, p. 640-645, 2017. <https://doi.org/10.1016/j.foodchem.2017.02.138>.

DARRIGUES, A.; HALL, J.; VAN DER KNAAP, E.; FRANCIS, D. M.; DUJMOVIC, N.; GRAY, S. Tomato analyzer-color test: a new tool for efficient digital phenotyping. **Journal of the American Society for Horticultural**

Science, v. 133, n. 4, p. 579-586, 2008.
<https://doi.org/10.21273/JASHS.133.4.579>.

DODIG, D.; BOŽINOVIC, S.; NIKOLIC, A.; ZORIC, M.; VANCETOVIC, J.; IGNJATOVIC-MICIC, D.; DELIC, N.; WEIGELT-FISCHER, K.; JUNKER, A.; ALTMANN, T. Image-derived traits related to mid-season growth performance of maize under nitrogen and water stress. **Frontiers in plant science**, v. 10, p. 814, 2019. <https://doi.org/10.3389/fpls.2019.00814>.

DOMINGOS, P. A few useful things to know about machine learning. **Communications of the ACM**, v. 55, n. 10, p. 78-87, 2012. DOI:10.1145/2347736.2347755

FAOSTAT-FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS (FAO). Participação dos continentes na produção de mandioca em 2019. FAOSTAT Database Gateway – FAO. Disponível em: <<http://www.fao.org/faostat/en/#data/QC>>. Acesso em: 18 Maio de 2021.

FERRARO, V.; PICCIRILLO, C.; TOMLINS, K.; PINTADO, M. E. Cassava (*Manihot esculenta* Crantz) and Yam (*Dioscorera* spp.) crops and their derived foodstuffs: Safety, security and nutritional value. **Critical reviews in food science and nutrition**, v. 56, n. 16, p. 2714-2727, 2016. <https://doi.org/10.1080/10408398.2014.922045>.

FIGÀS, M. R.; PROHENS, J.; RAIGÓN, M. D.; FERNÁNDEZ-DE-CÓRDOVA, P.; FITA, A.; SOLER, S. Characterization of a collection of local varieties of tomato (*Solanum lycopersicum* L.) using conventional descriptors and the high-throughput phenomics tool Tomato Analyzer. **Genetic Resources and Crop Evolution**, v. 62, n. 2, p. 189-204, 2015. <https://doi.org/10.1007/s10722-014-0142-1>.

FUHRMANN, E.; Vieira, E. A.; FIALHO, J. de F.; FALEIRO, F. G.; CARVALHO, L. J. C. B. Agronomic performance and biochemical attributes of yellow-pulped

elite sweet cassava clones. **Científica**, v. 47, n. 1, p. 77-82, 2019. <http://dx.doi.org/10.15361/1984-5529.2019v47n1p77-82>.

GONZÁLEZ, C. S.; RODRÍGUEZ-PULIDO, F. J.; STINCO, C. M.; TERRAB, A.; DÍAZ-MORENO, C.; FUENMAYOR, C.; HEREDIA, F. J. Carotenoid profile determination of bee pollen by advanced digital image analysis. **Computers and Electronics in Agriculture**, v. 175, p. 105601, 2020. <https://doi.org/10.1016/j.compag.2020.105601>.

HADIMANI, L.; MITTAL, N. Development of a computer vision system to estimate the colour indices of Kinnow mandarins. **Journal of food science and technology**, v. 56, n. 4, p. 2305-2311, 2019. <https://doi.org/10.1007/s13197-019-03641-9>.

IKEOGU, U. N.; AKDEMIR, D.; WOLFE, M. D.; OKEKE, U. G.; AMAEFULA, C.; JANNINK, J. L.; EGESI, C. N. Genetic Correlation, Genome-Wide Association and Genomic Prediction of Portable NIRS Predicted Carotenoids in Cassava Roots. **Frontiers in Plant Science**, v. 10, p. 1570, 2019. <https://doi.org/10.3389/fpls.2019.01570>.

IKEOGU, U. N.; DAVRIEUX, F.; DUFOUR, D.; CEBALLOS, H.; EGESI, C. N.; JANNINK, J. L. Rapid analyses of dry matter content and carotenoids in fresh cassava roots using a portable visible and near infrared spectrometer (Vis/NIRS). **PloS one**, v. 12, n. 12, p. 1-17, 2017. <https://doi.org/10.1371/journal.pone.0188918>.

ITLE, R. A.; KABELKA, E. A. Correlation between L* a* b* color space values and carotenoid content in pumpkins and squash (*Cucurbita* spp.). **HortScience**, v. 44, n. 3, p. 633-637, 2009. <https://doi.org/10.21273/HORTSCI.44.3.633>.

JARVIS, A.; RAMIREZ-VILLEGAS, J.; CAMPO, B. V. H.; NAVARRO-RACINES, C. Is cassava the answer to African climate change adaptation?. **Tropical Plant Biology**, v. 5, n. 1, p. 9-29, 2012. <https://doi.org/10.1007/s12042-012-9096-7>.

KASSAMBARA, A.; MUNDT, F. **factoextra: Extract and Visualize the Results of Multivariate Data Analyses**. R package version 1.0.5, 2017.

KLJAK, K.; GRBEŠA, D.; KAROLYI, D. Reflectance colorimetry as a simple method for estimating carotenoid content in maize grain. **Journal of Cereal Science**, v. 2, n. 59, p. 109-111, 2014. DOI: 10.1016/j.jcs.2013.12.004.

KUHN, M. Caret package. **Journal of statistical software**. 2008. Disponível em: <<http://caret.r-forge.r-project.org/>>. Acesso em: 02 de Junho de 2021.

LI, M.; FRANK, M. H.; MIGICOVSKY, Z. ColourQuant: a high-throughput technique to extract and quantify colour phenotypes from plant images. **arXiv 1903.01652**, 2019. <http://arxiv.org/abs/1903.01652>.

LIAKOS, K. G.; BUSATO, P.; MOSHOU, D.; PEARSON, S.; BOCHTIS, D. Machine learning in agriculture: A review. **Sensors**, v. 18, n. 8, p. 2674, 2018. <https://doi.org/10.3390/s18082674>.

MORESCO, R.; AFONSO, T.; UARROTA, V. G.; NAVARRO, B. B.; NUNES, E. D. C.; ROCHA, M.; MARASCHIN, M. Classification tools for carotenoid content estimation in *Manihot esculenta* via metabolomics and machine learning. In: **International Conference on Practical Applications of Computational Biology & Bioinformatics**. Springer, Cham, 2017. p. 280-288. https://doi.org/10.1007/978-3-319-60816-7_34

MORESCO, R.; UARROTA, V. G.; PEREIRA, A.; TOMAZZOLI, M.; NUNES, E. D. C.; PERUCH, L. A. M.; GAZZOLA, J.; COSTA, C.; ROCHA, M.; MARASCHIN, M. UV-visible scanning spectrophotometry and chemometric analysis as tools for carotenoids analysis in cassava genotypes (*Manihot esculenta* Crantz). **Journal of integrative bioinformatics**, v. 12, n. 4, p. 27-38, 2015. doi: 10.2390/biecoll-jib-2015-28.

NANKAR, A. N.; TRINGOVSKA, I.; GROZEVA, S.; TODOROVA, V.; KOSTOVA, D. Application of high-throughput phenotyping tool Tomato Analyzer to characterize Balkan *Capsicum* fruit diversity. **Scientia Horticulturae**, v. 260, p. 108862, 2020. <https://doi.org/10.1016/j.scienta.2019.108862>.

PACE, B. CEFOLA, M.; RENNA, F.; RENNA, M.; SERIO, F.; ATTOLICO, G. Multiple regression models and computer vision systems to predict antioxidant activity and total phenols in pigmented carrots. **Journal of food engineering**, v. 117, n. 1, p. 74-81, 2013. <https://doi.org/10.1016/j.jfoodeng.2013.02.005>.

PARMAR, A.; STURM, B.; HENSEL, O. Crops that feed the world: Production and improvement of cassava for food, feed, and industrial uses. **Food Security**, v. 9, n. 5, p. 907-927, 2017. <https://doi.org/10.1007/s12571-017-0717-8>.

PATHARE, P. B.; OPARA, U. L.; AL-SAID, F. A-J. Colour measurement and analysis in fresh and processed foods: a review. **Food and bioprocess technology**, v. 6, n. 1, p. 36-60, 2013. <https://doi.org/10.1007/s11947-012-0867-9>.

PEPRAH, B. B.; PARKES, E. Y.; HARRISON, O. A.; VAN BILJON, A.; STEINER-ASIEDU, M.; LABUSCHAGNE, M. T. Proximate composition, cyanide content, and carotenoid retention after boiling of provitamin A-rich cassava grown in Ghana. **Foods**, v. 9, n. 12, p. 1800, 2020. <https://doi.org/10.3390/foods9121800>.

R CORE TEAM. 2021. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Disponível em: <<https://www.R-project.org/>>

RODRIGUEZ-AMAYA, D. B.; KIMURA, M. **Harvest Plus handbook for carotenoid analysis**. Washington: IFPRI, 2004. 58p.

SALTZMAN, A.; BIROL, E.; OPARINDE, A.; ANDERSSON, M. S.; ASARE-MARFO, D.; DIRESSIE, M. T.; GONZALEZ, C.; LIVIDINI, K.; MOURSI, M.; ZELLER, M. (Availability, production, and consumption of crops biofortified by plant breeding: current evidence and future potential. **Annals of the New York Academy of Sciences**, v. 1390, n. 1, p. 104-114, 2017. DOI: 10.1111/nyas.13314.

SANAEIFAR, A.; JAFARI, A.; GOLMAKANI, M-T. Fusion of dielectric spectroscopy and computer vision for quality characterization of olive oil during storage. **Computers and Electronics in Agriculture**, v. 145, p. 142-152, 2018. DOI: <https://doi.org/10.1016/j.compag.2017.12.035>.

SÁNCHEZ, T.; CEBALLOS, H.; DUFOUR, D.; ORTIZ, D.; MORANTE, N.; CALLE, F.; FELDE, T. Z.; DOMÍNGUEZ, M.; DAVRIEUX, F. Prediction of carotenoids, cyanide and dry matter contents in fresh cassava root using NIRS and Hunter color techniques. **Food Chemistry**, v. 151, p. 444–451, 2014. <https://doi.org/10.1016/j.foodchem.2013.11.081>.

SAYRE, R.; BEECHING, J. R.; CAHOON, E. B.; EGESI, C.; FAUQUET, C.; FELLMAN, J.; FREGENE, M.; GRUISSEM, W.; MALLOWA, S.; MANARY, M.; MAZIYA-DIXON, B.; MBANASO, A.; SCHACHTMAN, D. P.; SIRITUNGA, D.; TAYLOR, N.; VANDERSCHUREN, H.; ZHANG, P. The BioCassava plus program: biofortification of cassava for sub-Saharan Africa. **Annual review of plant biology**, v. 62, p. 251-272, 2011. <https://doi.org/10.1146/annurev-arplant-042110-103751>.

SELVARAJ, M. G.; Valderrama, M.; Guzman, D.; Valencia, M.; Ruiz, H.; Acharjee, A. Machine learning for high-throughput field phenotyping and image processing provides insight into the association of above and below-ground

traits in cassava (*Manihot esculenta* Crantz). **Plant methods**, v. 16, n. 1, p. 1-19, 2020. <https://doi.org/10.1186/s13007-020-00625-1>.

SINGH, Y.; BHATIA, P. K.; SANGWAN, O. A review of studies on machine learning techniques. **International Journal of Computer Science and Security**, v. 1, n. 1, p. 70-84, 2007.

SOUZA, L. D.; SOUZA, L. S.; GOMES, J. C. (2006) Exigências edáficas da cultura da mandioca. In: SOUZA, L. S.; FARIAS, A. R. N.; MATTOS, P. L. P.; FUKUDA, W. M. G. (Eds.), **Aspectos Socioeconômicos e Agronômicos da Mandioca**. Cruz Das Almas: Embrapa Mandioca e Fruticultura Tropical, p. 70–214.

STINCO, C. M.; RODRÍGUEZ-PULIDO, F. J.; ESCUDERO-GILETE, M. L.; GORDILLO, B.; VICARIO, I. M.; MELÉNDEZ-MARTÍNEZ, A. J. Lycopene isomers in fresh and processed tomato products: Correlations with instrumental color measurements by digital image analysis and spectroradiometry. **Food research international**, v. 50, n. 1, p. 111-120, 2013. <https://doi.org/10.1016/j.foodres.2012.10.011>.

WALTER, J.; EDWARDS, J.; CAI, J.; MCDONALD, G.; MIKLAVCIC, S. J.; KUCHEL, H. High-throughput field imaging and basic image analysis in a wheat breeding programme. **Frontiers in plant science**, v. 10, p. 449, 2019. <https://doi.org/10.3389/fpls.2019.00449>.

WAN, P.; TOUDESCHI, A.; TAN, H.; EHSANI, R. A methodology for fresh tomato maturity detection using computer vision. **Computers and electronics in agriculture**, v. 146, p. 43-50, 2018. <https://doi.org/10.1016/j.compag.2018.01.011>.

WANG, S.; SUMMERS, R. M. Machine learning and radiology. **Medical image analysis**, v. 16, n. 5, p. 933-951, 2012. <https://doi.org/10.1016/j.media.2012.02.005>.

WASONGA, D. O.; YAW, A.; KLEEMOLA, J.; ALAKUKKU, L.; MÄKELÄ, P. S. Red-green-blue and multispectral imaging as potential tools for estimating growth and nutritional performance of cassava under deficit irrigation and potassium fertigation. **Remote Sensing**, v. 13, n. 4, p. 598, 2021. <https://doi.org/10.3390/rs13040598>.

YE, X.; IZAWA, T.; ZHANG, S. Rapid determination of lycopene content and fruit grading in tomatoes using a smart device camera. **Cogent Engineering**, v. 5, n. 1, p. 1504499, 2018. <https://doi.org/10.1080/23311916.2018.1504499>.

ZAOUAY, F.; MENA, P.; GARCIA-VIGUERA, C.; MARS, M. Antioxidant activity and physico-chemical properties of Tunisian grown pomegranate (*Punica granatum* L.) cultivars. **Industrial Crops and Products**, v. 40, p. 81-89, 2012. <https://doi.org/10.1016/j.indcrop.2012.02.045>.