

ANÁLISE QUANTITATIVA DE GRÃOS DE ARROZ

Ferreira, Alex Naves¹, Martins, Ana Carolina Faria¹, Cardoso, Felipe Pereira¹, Miranda, Rafael Novais¹, Martins, Vinicius Samuel¹

¹Departamento de Biologia, Universidade Federal de Lavras, Lavras-MG, CEP 37.200-000, Brasil

Resumo

O arroz, *Oryza sativa* L., pertence à família das gramíneas, taxonomicamente denominado de Poaceae. É uma das culturas anuais de maior importância social e econômica no Brasil. Através do melhoramento genético é possível obter grãos com características favoráveis. Neste contexto, o objetivo deste trabalho foi avaliar vinte diferentes linhagens quanto as dimensões de largura, comprimento, área e razão. As avaliações foram realizadas utilizando imagens obtidas utilizando o equipamento Groundeye e processamento de imagens via Pycharm. Através da mensuração da largura e comprimento foi possível observar que todos os grãos possuem morfologia longa, típica preferência do consumidor. Além disso foi possível observar uma correlação positiva entre área de grãos e peso de 1000 grãos, o que viabiliza a seleção indireta utilizando a área.

Palavras-chave: *Oryza sativa*, processamento de imagem,

1. INTRODUÇÃO

O arroz pertence a família Poaceae (gramíneas) e ao gênero *Oryza*, a espécie *Oryza sativa* L. é a mais cultivada no Brasil entre as vinte espécies existente (JULIANO, 1993). É uma das culturas anuais de maior importância social e econômica no Brasil, visto que, juntamente com o feijão, é o principal componente da dieta diária dos brasileiros (WALTER; MARCHEZAN; AVILA, 2008). O arroz possui carboidratos, proteínas, lipídios, vitaminas (B1, B2 e B3) e minerais, entretanto a quantidade de nutrientes e minerais presente no grão e influenciados por fatores como qualidade do solo, fertilizantes, condições climáticas, beneficiamento/processamento, armazenamento, entre outros (CONAB, 2015).

O grande consumo populacional por esse alimento fez com que o arroz seja constantemente objeto de diversos estudos (MAIONE; BARBOSA, 2019). Ele é um dos

principais cereais consumidos e produzidos no mundo, caracterizando-se como principal alimento para mais de 50% da população mundial. A sua importância é destacada, consequentemente, em países em desenvolvimento, tais como o Brasil, desempenhando papel estratégico em níveis econômico e social (FAO, 2018)

O aumento exponencial na população mundial, consequentemente, também na demanda por alimentos, fez com que a produção de arroz mundial aumentasse. Esse acréscimo na produção mundial de arroz, foi de 285 milhões de toneladas em 1961 para 741 milhões de toneladas em 2016 (FAO, 2018). Ele é tradicionalmente, um dos produtos alimentícios mais consumidos no Brasil, sendo o seu consumo anual estimado em 12,07 milhões de toneladas, na safra 17/18 e seu consumo per capita é de 34,5 kg por ano (CONAB, 2017; FAO, 2015).

Análises de imagem do arroz tem sido amplamente realizada na literatura recente com o auxílio de análises multivariadas de dados e técnicas de fenotipagem por imagem. Ambos os processos oferecem métodos poderosos capazes de realizar análises estatísticas e análises de imagem sobre dados e imagens de arroz.

Neste trabalho, o objeto foi avaliar vinte diferentes linhagens quanto as dimensões de largura, comprimento, área e razão. Para isto utilizamos o utilitário Python Software Foundation License, para realizarmos a avaliação das imagens gerados pelo equipamento GroundEye[®].

2. MATERIAL E MÉTODOS

2.1 Origem das linhagens

Neste trabalho foram utilizadas 20 linhagens de VCU de arroz (Tabela 1) com três repetições obtidas da safra 2018/2019. Os grãos estão depositados na Universidade Federal de Lavras, Departamento de Agronomia (DAG).

Tabela 1: Linhagens de arroz oriundas do VCU.

Tratamento	Linhagem	Tratamento	Linhagem
1	CMG F6 LAM 20-2	11	CMG ERF 85-15
2	CMG ERF 85-14	12	CMG ERF 85-3
3	BRS Esmeralda	13	CMG F6 LAV 1-7
4	CMG ERF 221-16	14	CMG ERF 85-4
5	CMG 2119	15	CMG ERF 222-1
6	BRSMG Caçula	16	CMG ERF 46-1
7	CMG ERF 85-6	17	CMG ERF 221-19
8	CMG ERF 81-2	18	CMG ERF 221-29
9	CMG ERF 81-6	19	CMG ERF 85-13
10	CMG 2085	20	Multilinha

2.2 Morfologia dos grãos

A dimensão do grão de arroz foi obtida a partir de análise de imagens. Foi feita uma subamostra de 100 grãos oriundos de cada parcela e em seguida realizadas as análises. A captura das imagens dos grãos foi efetuada utilizando o equipamento Groundeye, o qual foi desenvolvido pela Tbit Tecnologia e Sistemas. Os grãos foram dispostos em a bandeja de vidro do equipamento sem nenhuma posição definida e posteriormente as imagens foram capturadas por câmeras de alta resolução situadas no interior do equipamento. As configurações do Groundeye foram processadas para calibração da cor do fundo, luminosidade, e dimensões do grão. Após este processo, foram realizadas as análises biométricas dos grãos, nas quais foram obtidos os valores de comprimento e largura dos grãos. Os grãos foram classificados de acordo com os padrões definidos.

Os grãos foram classificados de acordo com os padrões definidos por Brasil (2012): longo-fino ($C \geq 6$ mm; $L \leq 2,17$ mm; $C/L > 2,75$), longo ($C \geq 6$ mm), médio ($C < 6$ e ≥ 5 mm) e curto ($C < 5$ mm).

2.3 Processamento das imagens

Foram utilizados computador tipo laptop e os softwares Python e Pycharm bem como as bibliotecas *opencv*, *numpy*, *pandas* e *Scikit-image*. A aquisição das imagens foi feita em RGB, no entanto ao se carregar as imagens no programa com o pacote *opencv* o

sistema de cor da imagem era altera para BGR, devido a este problema as imagens as imagens tiveram que ser transformadas novamente para RGB. Com a imagem em RGB foi realizado a segmentação do sistema de cor nos três canais (R, G, B), posteriormente foi realizado o *thresholding* (Figura 1), para esta tarefa foi utilizado o método do limiar de Otsu automático. Na técnica de Otsu foi utilizado o canal R onde foi aplicado o filtro bilateral (proporciona uma pequena perda de foco/nitidez da imagem em relação a imagem original). Ao final se obteve uma imagem segmentada (imagem com o fundo preto e os grãos colorido) e uma binária (imagem branca e preta).

Na próxima etapa foi utilizado o reconhecimento dos contornos dos grãos para identifica os contornos de cada objeto (grão) que está na imagem. Com os dados dos contornos foi utilizado o pacote *Scikit-image* para obter dos dados de eixo maior, eixo menor, área, razão e número de grãos, estes dados foram salvos em uma tabela utilizando o pacote *pandas*.

O que está descrito a cima foi aplicado em todas as imagens, aquelas imagens onde o programa não rodou foi realizado modificações no script precisamente alterando os canais, além do sistema de cor RGB foi utilizado o Lab e YCrCb destes sistemas de cor foi aplicado o canal que melhor se adaptou a imagem.

2.4 Análise de correlação entre a área e peso de grãos.

O peso de 1000 grãos das linhagens foi obtido usando a balança de precisão Bel 0,01g, 2200g S2202H. A partir dos dados de peso foi realizado uma análise de variância para que se pudesse inferir sobre a correlação entre peso dos grãos e a área. A área foi obtida utilizando as fotos em que a metodologia foi descrita no item anterior.

3. RESULTADOS E DISCUSSÃO

A limiarização da imagem pelo método de OTSU propiciou bons resultados, oferecendo boa obtenção de contorno entre os grãos amostrados (Figura 1). Diante disso, foi possível a obtenção das características referentes a cada grão, como: largura, comprimento e a área (Tabela 2).

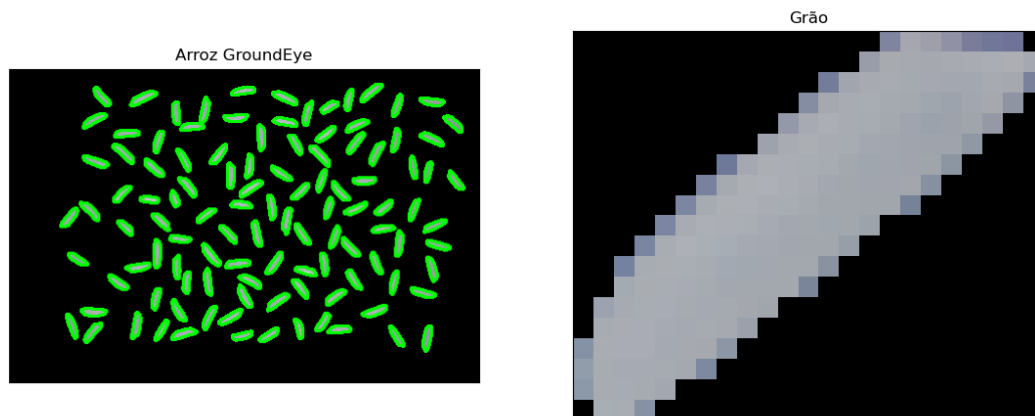


Figura 1: Segmentação da imagem com seus respectivos contornos.

Tabela 2: Média da largura, comprimento, razão, área e peso dos grãos de arroz para cada linhagem.

Tratamento	Largura	Comprimento	Razão	Área	MMG (g)
1	8.57	23.84	2.81	128.9	24.27
2	9.59	24.95	2.64	153.36	30.65
3	8.15	24.13	3.02	122.13	23.64
4	8.21	25.89	3.17	136.84	28.24
5	7.99	22.92	3	112.67	23.12
6	8.07	24.18	3.06	126.01	27.05
7	10.2	27.56	3.13	132.66	28.79
8	9.83	24.9	2.88	126.03	28.25
9	8.38	24.64	2.97	130.56	26.96
10	8.76	26.34	3.11	140.4	29.98
11	8.68	27	3.21	142.58	31.16
12	8.62	27.17	3.23	144.14	31.01
13	7.29	24.94	3.44	116.52	24.32
14	8.12	25.76	3.19	135.48	28.18
15	8.50	22.46	2.67	118.98	22.26
16	7.74	24.15	3.20	112.85	24.12
17	9.03	27.45	3.18	143.83	31.59
18	8.40	26.43	3.31	127.12	27.43
19	8.47	23.75	2.84	126.35	26.71
20	8.68	25.87	3.10	132.79	28.39

É possível observar que as características área do grão e massa de mil grãos (MMG) seguem uma tendência linear positiva, esse resultado poderia tornar viável a seleção indireta da MMG por análise de imagem, o que seria mais simples e eficiente

(FIGURA 2). Para ambas características, os genótipos CMG ERF 221-19, CMG ERF 85-14, CMG ERF 85-15 e CMG ERF 85-3 foram promissores.

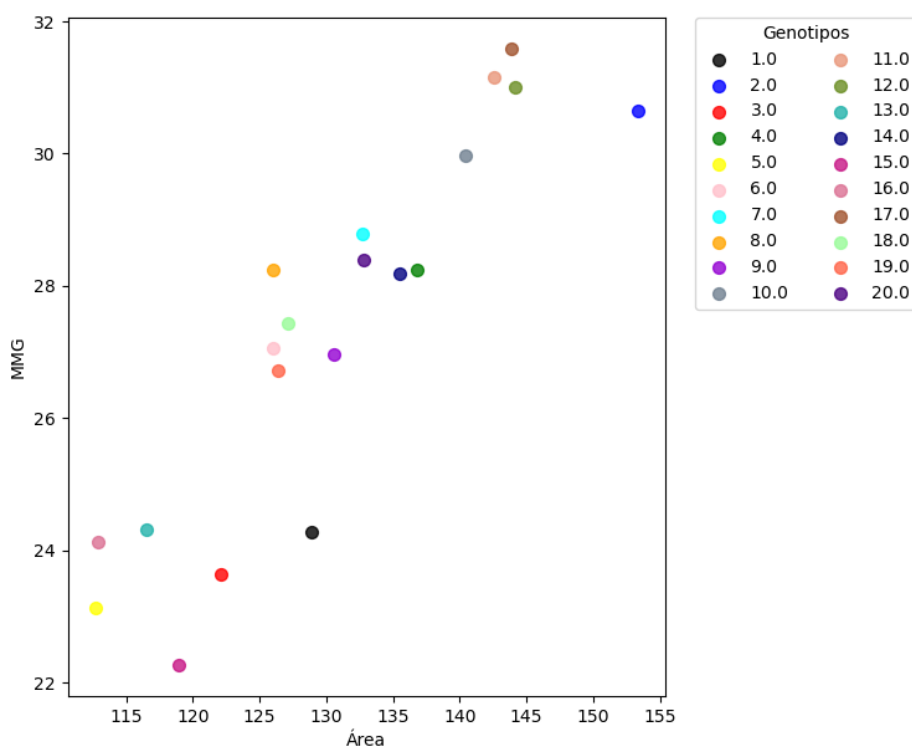


Figura 2: Gráfico de dispersão das médias dos genótipos para as características massa de mil grãos (MMG) e área do grão.

Todos os grãos foram classificados como longos de acordo com os padrões estabelecidos por Brasil (2012), visto que todos possuem o comprimento maior que 6mm. Tal fato pode ser explicado pela seleção de linhagens com morfologia longa dos grãos, já que este tipo de morfologia é mais consumido pela população.

3.1 Análise de correlação entre a área e peso de grãos.

Houve uma correlação positiva entre a massa de 1000 grãos e a área (estimada por meio dos softwares Python e Pycharm). Foram observados correlação genotípica em sentido favorável ($r = 0,94$) de alta magnitude entre as variáveis área e peso de 1000 grãos (Anexo). Pode-se inferir, que à medida que aumenta à área dos grãos ocorre incremento na massa de 1000 grãos. Há evidências, que ao realizar a seleção de linhagens por imagem (seleção indireta) com maior área dos grãos, ocorre incremento em massa de 1000 grãos.

A vantagem ao praticar seleção indireta é obter ganhos genéticos em caracteres de difícil mensuração ou que apresentam herdabilidade baixa. O caráter massa de 1000 grãos apresenta herdabilidade de alta magnitude de acordo com saída do programa genes em anexo. Contudo, o mesmo caráter é controlado por muitos genes e apresenta de média a alta dificuldade de obtenção dos dados, sendo uma prática onerosa e trabalhosa. Ao realizar seleção indireta, o melhorista aumenta suas chances de obter sucesso. Isto é, alcançar em menor tempo a melhor linhagem ou cultivar.

4. CONCLUSÕES

É possível obter métricas do formato de grãos de arroz utilizando análise de imagens. Portanto, é verificada diferença significativas entre as linhagens em relação ao caráter área de grãos. Desta forma, a análise de imagens é uma ferramenta poderosa que nos auxilia na prática de seleção indireta para selecionar linhagens com maior massa de mil grãos.

5. REFERÊNCIAS

BRASIL. Ministério da Agricultura, Pecuária e Abastecimento. Instrução Normativa nº 02, de 6 de fevereiro de 2012. Disponível em: <<http://sistemasweb.agricultura.gov.br/sislegis/action/detalhaAto.do?method=visualizarAtoPortalMapa&chave=918108049>>. Acesso em: 22 nov. 2017

CONAB. COMPANHIA NACIONAL DE ABASTECIMENTO. A cultura do arroz. Brasília: Conab, p. 180, 2015. Disponível em: <<http://www.conab.gov.br>>. Acesso em: 01/08/2020.

FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS (FAO). OECD-FAO Agricultural Outlook 2015-2024. Paris: OECD-FAO: 2015. 143 p.

FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS (FAO). OECD-FAO Agricultural Outlook 2015-2024. Paris: OECD-FAO: 2018. 101 p.

JULIANO, B. O. Rice in human nutrition. Rome. FAO, 1993. Disponível em: <<http://www.fao.org/docrep/t0567e/T0567E00.htm#Contents>>. Acesso em: 30 agosto. 2020.

MAIONE, Camila; BARBOSA, Rommel Melgaço. Aplicações recentes de métodos de análise multivariada de dados na autenticação de arroz e os parâmetros mais analisados: Uma revisão. Revisões críticas em ciência alimentar e nutrição, v. 59, n. 12, pág. 1868-1879, 2019.

WALTER, M., MARCHEZAN, E., AVILA, L. A. Arroz: composição e características nutricionais. Ciência Rural, Santa Maria, v.38, n.4, p. 1184-1192, 2008.

ANEXO A

```

=====
Programa GENES                                ANOVA EM BLOCOS AO ACASO
Arquivo                                         de                      dados
C:\Users\Cliente\Documents\GitHub\visao_computacional_Vinicius\Trabalho
Final\dados correlação.txt
Número de variáveis                           2
Número de genótipos                          20
Número de repetições                         3
Data                                           08-31-2020
=====

```

ANÁLISE DA VARIÁVEL => AREA

Correlação entre blocos

Blocos	Covariância	Correlação
1 x 2	121.187128	.914507
1 x 3	118.41138	.92901
2 x 3	104.989439	.936454
Média	114.862649	.926657

ANÁLISE DE VARIÂNCIA DA VARIÁVEL => AREA

FV	GL	SQ	QM	F
Probabilidade (%)				
BLOCOS	2	8.17143	4.085715	
TRATAMENTOS	19	6740.112024	354.742738	34.9335
RESÍDUO	38	385.882027	10.15479	.0**
TOTAL	59	7134.165481		
MÉDIA		130.213297	CV(%)	2.44726
MÍNIMO		101.5	MÁXIMO	153.75
DMS-Tukey (1%)		11.483408	DMS-Tukey (5%)	9.896197

Estimativas de Parâmetros

VARIÂNCIA FENOTÍPICA (média)	118.247579
VARIÂNCIA AMBIENTAL (média)	3.38493
VARIÂNCIA GENOTÍPICA (média)	114.862649
HERDABILIDADE (US: média da família) - %	97.1374
CORRELAÇÃO INTRACLASSE (US: parcela)- %	91.8773
COEFICIENTE DE VARIAÇÃO GENÉTICO (%)	8.2306
RAZÃO CVg/CVe	3.3632
$r = (1 - 1/F)^{0.5}$.9856

ANÁLISE DA VARIÁVEL => MMG

Correlação entre blocos

Blocos	Covariância	Correlação
--------	-------------	------------

1	x	2	8.281921	.97089
1	x	3	7.790768	.95867
2	x	3	7.255021	.941749
Média			7.775904	.957103

—

ANÁLISE DE VARIÂNCIA DA VARIÁVEL => MMG

FV	GL	SQ	QM	F
Probabilidade(%)				
BLOCOS	2	4.32052	2.16026	
TRATAMENTOS	19	450.170563	23.693188	64.8306
RESÍDUO	38	13.8876	.365463	.0**
TOTAL	59	468.3787		

MÉDIA	27.374417	CV(%)	2.208396
MÍNIMO	22.1875	MÁXIMO	31.8375
DMS-Tukey(1%)	2.178498	DMS-Tukey(5%)	1.877391

Estimativas de Parâmetros

VARIÂNCIA FENOTÍPICA (média)	7.897729
VARIÂNCIA AMBIENTAL (média)	.121821
VARIÂNCIA GENOTÍPICA (média)	7.775908
HERDABILIDADE (US: média da família) - %	98.4575
CORRELAÇÃO INTRACLASSE (US: parcela)- %	95.511
COEFICIENTE DE VARIAÇÃO GENÉTICO (%)	10.1866
RAZÃO CVg/CVe	4.6127
$r = (1 - 1/F)^{0.5}$.9923

COVARIÂNCIAS FENOTÍPICAS

118.247579 28.075309
28.075309 7.897729

COVARIÂNCIAS GENOTÍPICAS

114.862649 28.036706
28.036706 7.775908

COVARIÂNCIAS RESIDUAIS

10.15479 .115807
.115807 .365463

CORRELAÇÕES FENOTÍPICAS

1. .9187
.9187 1.

CORRELAÇÕES GENOTÍPICAS	
1. .9381	
.9381 1.	
CORRELAÇÕES RESIDUAIS	
1. .0601	
.0601 1.	
MÉDIAS DAS VARIÁVEIS	
130.6366	24.7667
151.5362	30.9917
119.7051	23.9792
135.9053	28.0625
113.3979	23.1042
125.225	27.4458
131.9581	29.3042
129.1819	27.8375
130.1944	26.7958
139.8077	30.2792
142.6625	30.875
142.7924	30.5958
115.4503	24.3667
135.5073	28.4133
117.6813	22.375
112.5361	24.175
143.8286	31.5875
127.1221	27.4333
126.3473	26.7083
132.79	28.3917