

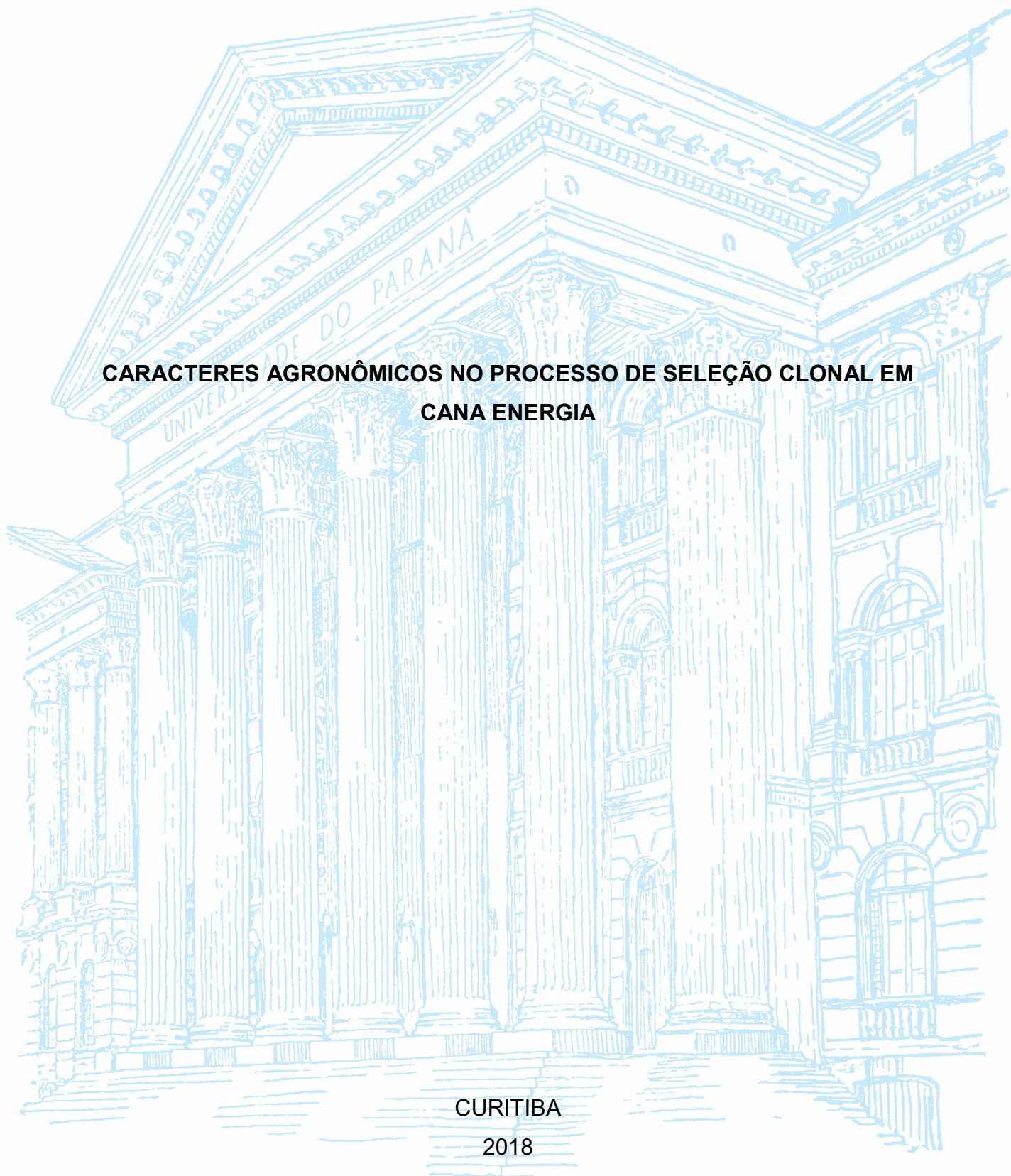
UNIVERSIDADE FEDERAL DO PARANÁ

JULIANE BORELLA

**CARACTERES AGRONÔMICOS NO PROCESSO DE SELEÇÃO CLONAL EM  
CANA ENERGIA**

CURITIBA

2018



JULIANE BORELLA

**CARACTERES AGRONÔMICOS NO PROCESSO DE SELEÇÃO CLONAL EM  
CANA ENERGIA**

Dissertação apresentada ao Programa de Pós-Graduação em Agronomia, Área de Concentração em Produção Vegetal, Departamento de Fitotecnia e Fitossanitarismo, Setor de Ciências Agrárias, Universidade Federal do Paraná, como parte das exigências para obtenção do título de Mestre em Ciências.

Orientador: Prof. Dr. João Carlos Bepalhok Filho

Co-orientador: Prof. Dr. Bruno Portela Brasileiro

CURITIBA

2018

FICHA CATALOGRÁFICA ELABORADA PELO SISTEMA DE BIBLIOTECAS/UFPR -  
BIBLIOTECA DE CIÊNCIAS AGRÁRIAS, DOUGLAS ALEX JANKOSKI CRB 9/1167  
COM OS DADOS FORNECIDOS PELO(A) AUTOR(A)

B731c Borella, Juliane  
Caracteres agronômicos no processo de seleção clonal  
em cana energia / Juliane Borella. - Curitiba, 2018.  
70 f. : il., grafs., tabs.

Dissertação (Mestrado) - Universidade Federal do  
Paraná. Setor de Ciências Agrárias, Programa de Pós-  
Graduação em Agronomia - (Produção Vegetal).  
Orientador: João Carlos Bepalhok Filho  
Coorientador: Bruno Portela Brasileiro

1. Cana-de-açúcar - Melhoramento genético. 2. Biomassa  
vegetal. 3. Genética vegetal. I. Bepalhok Filho, João Carlos.  
II. Brasileiro, Bruno Portela. III. Título. IV. Universidade  
Federal do Paraná.

CDU 633.61



MINISTÉRIO DA EDUCAÇÃO  
SETOR SETOR DE CIÊNCIAS AGRARIAS  
UNIVERSIDADE FEDERAL DO PARANÁ  
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO  
PROGRAMA DE PÓS-GRADUAÇÃO AGRONOMIA  
(PRODUÇÃO VEGETAL)

## TERMO DE APROVAÇÃO


Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em AGRONOMIA (PRODUÇÃO VEGETAL) da Universidade Federal do Paraná foram convocados para realizar a arguição da dissertação de Mestrado de **JULIANE BORELLA** intitulada: **CARACTERES AGRONÔMICOS NO PROCESSO DE SELEÇÃO CLONAL EM CANA ENERGIA**, após terem inquirido a aluna e realizado a avaliação do trabalho, são de parecer pela sua           APROVAÇÃO           no rito de defesa.

A outorga do título de mestre está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

Curitiba, 22 de Fevereiro de 2018.

  
JOÃO CARLOS BESPALHOK FILHO  
Presidente da Banca Examinadora (UFPR)

  
BRUNO PORTELA BRASILEIRO  
Avaliador Interno (UFPR)

  
EDSON PEREZ GUERRA  
Avaliador Externo (UNICENTRO)

*A Deus,*

*A minha mãe Elenir Inês Borella e ao*

*Meu pai Caitano Borella*

*Dedico*

*“Conheça todas as teorias, domine todas as técnicas, mas ao tocar uma alma humana, seja apenas outra alma humana.”*

*Carl Jung*

## **AGRADECIMENTOS**

Ao Professor João Carlos Bessalho Filho pela orientação, amizade, paciência em orientar e constante motivação para alcançar este objetivo. Obrigada pelo voto de confiança depositado em agosto de 2015 quando me aceitou para realizar o estágio de conclusão do curso de Agronomia.

Ao Professor Bruno Portela Brasileiro, pela Co-orientação, disposição e auxílio para realização do trabalho.

Ao professor Alexandre Behling, que sempre se mostrou disponível para sanar as dúvidas, obrigada pelo auxílio e amizade.

A minha família Elenir, Caitano e Junior Borella por me mostrarem o caminho do bem, por apoiarem nas minhas decisões e me ensinarem que o caráter e o respeito são nossos melhores aliados.

Ao Jonathan William Trautenmuller, pelo companheirismo nesses sete anos, obrigada pelos ensinamentos e paciência.

Ao Programa de Pós-graduação em Agronomia, Área de Concentração em Produção Vegetal, da Universidade Federal do Paraná pela oportunidade de realizar os estudos necessários para a conclusão do Mestrado.

Aos técnicos do Programa de Melhoramento Genético de Cana-de-açúcar da Universidade Federal do Paraná (PMGCA/UFPR/RIDES), pelo apoio na implantação e coleta de dados do experimento, além da amizade e consideração.

A CAPES, pela concessão da bolsa no período de realização do mestrado.

A todos os professores da Pós-Graduação que contribuíram com a minha qualificação profissional, aos amigos do curso pelo companheirismo durante as atividades acadêmicas. Aos técnicos administrativos e funcionários da UFPR.

Aos amigos Mariana Almeida Franca, Tales Romano, Gildemberg Amorin Leal Junior, pelas boas discussões e convivência durante este caminho. E aos demais que fizeram parte dessa caminhada meu agradecimento.

## RESUMO

O conhecimento dos caracteres envolvidos no processo de seleção clonal, nas diferentes etapas do melhoramento genético, bem como a utilização de estratégias que contribuam no momento da seleção, são importantes para auxiliar na seleção dos novos genótipos de cana energia. O objetivo desse trabalho foi determinar a importância dos caracteres agrônômicos envolvidos no processo de seleção de clones, visando otimizar o processo de seleção em populações destinadas a obtenção de cana energia. O experimento foi realizado na Estação Experimental de Paranaíba, pertencente à Universidade Federal do Paraná. Anteriormente, 220 clones selecionados na primeira fase de teste (T1), foram plantados em junho de 2014 em delineamento de blocos incompletos. As parcelas foram constituídas de duas linhas de 5,0 metros espaçadas 1,4 metros, com plantio de 18 gemas por metro. As avaliações dos componentes de produção e de qualidade foram realizadas nos ciclos de cana-planta e primeira soqueira. Foi feito no ciclo de primeira soqueira a atribuição de notas para os caracteres despalha, diâmetro do colmo, altura do colmo, número de colmos, presença de brotação lateral, saliência da gema, hábito de crescimento, florescimento, chochamento, nota para vigor, incidência de carvão, ferrugem marrom e ferrugem alaranjada. Os dados obtidos foram analisados por meio da análise de trilha e regressão logística. Nos ciclos de cana-planta e primeira soqueira foram utilizados como caracteres dependentes, a tonelada de colmos por hectare com palha, tonelada de colmos por hectare sem palha, tonelada de fibra por hectare com palha e tonelada de sacarose por hectare sem palha. Para o ciclo de cana-planta os caracteres de maior efeito direto na tonelada de colmos por hectare com palha, tonelada de colmos por hectare sem palha e tonelada de fibra por hectare com palha foram peso médio do colmo e número de colmos. Já os de maior efeito direto na tonelada de sacarose por hectare sem palha foram peso médio do colmo, teor de sacarose aparente seguido de número de colmos. No ciclo de primeira soqueira, esses caracteres se mantiveram como os de maior efeito. Foram selecionados 38 clones pelo método massal na fase T2, enquanto nos modelos ajustados por regressão logística, denominados de qualitativo e quantitativo, selecionaram 42 e 22 clones, respectivamente. A capacidade preditiva dos modelos, qualitativo e quantitativo foi de 94% e 87%, respectivamente. A utilização de um modelo logístico composto de variáveis categóricas se mostrou eficiente, podendo ser utilizado no processo de seleção clonal em populações de cana energia.

Palavras-chave: melhoramento genético, biomassa, análise de trilha, regressão logística.

## ABSTRACT

The knowledge of the traits involved in the selection process of the clones, in the different stages of genetic improvement, as well as the use of strategies that contribute to the selection process, are important to assist in the selection of the new sugar cane genotypes. The objective of this work was to determine the importance of the agronomic traits involved in the selection process of clones, in order to optimize the selection process in populations destined to obtain energy cane. The experiment was carried out at the Paranavaí Experimental Station, belonging to the Federal University of Paraná. Previously, 220 clones selected in the first test phase (T1) were planted in June 2014 in an incomplete block design. The plots were constituted of two rows of 5.0 meters spaced 1.4 meters, with planting of 18 buds per meter. The evaluations of the production and quality components were carried out in the plant cane plant and first ratoon cycles. It was done in the cycle of plant cane the assignment of notes for the traits, straw removal, stalk diameter, stalk number, stalk height, sprouting lateral, growth habit, flowering, stalk without broth, note for plant vigor, incidence of smut, brown rust and orange rust. The data obtained were analyzed through path analysis and logistic regression. In the plant cane and first ratoon cycles, the ton of stalks per hectare with straw, ton of stalks per hectare without straw, ton of fiber per hectare with straw and ton of sucrose per hectare without straw were used as dependent characters. For the plant cane cycle, the characters with the greatest direct effect in the ton of stems per hectare with straw, ton of stems per hectare without straw, and ton of fiber per hectare with straw were average stem weight and number of stems. Those with the highest direct effect in the ton of sucrose per hectare without straw were mean weight of the stem, apparent sucrose content followed by stalk number. In the first ratoon cycle, these characters remained the most effective. The mass method selected 38 clones in T2, while logistic regression models, called qualitative and quantitative, selected 42 and 22 clones respectively. The predictive capacity of the models, qualitative and quantitative were 94% and 87%, respectively. The use of a logistic model composed of categorical variables was efficient and could be used in the process of clonal selection in energy cane populations.

Keywords: genetic improvement, biomass, path analysis, logistic regression.

## LISTA DE TABELAS

- TABELA 3.1 - ESTATÍSTICA DESCRITIVA DOS CARACTERES COMPONENTES DE PRODUÇÃO E DE QUALIDADE EM CANA ENERGIA NOS CICLOS DE CANA-PLANTA E PRIMEIRA SOQUEIRA..... 35
- TABELA 3.2 - ESTIMATIVAS DOS COEFICIENTES DE CORRELAÇÃO GENÉTICA ENTRE OS CARACTERES. OS VALORES DE CORRELAÇÃO ACIMA DA DIAGONAL PRINCIPAL SÃO REFERENTES AO CICLO DE CANA-PLANTA E ABAIXO DA DIAGONAL PRINCIPAL AO CICLO DE PRIMEIRA SOQUEIRA. .... 36
- TABELA 3.3 - EFEITOS DIRETOS E INDIRETOS DOS CARACTERES EXPLICATIVOS SOBRE OS CARACTERES PRINCIPAIS TONELADA DE COLMOS POR HECTARE COM PALHA (TCHcp) E TONELADA DE COLMO POR HECTARE SEM PALHA (TCHsp) EM CANA-PLANTA E PRIMEIRA SOQUEIRA NA SEGUNDA FASE DE TESTE..... 38
- TABELA 3.4 - EFEITOS DIRETOS E INDIRETOS DOS CARACTERES EXPLICATIVOS SOBRE OS CARACTERES PRINCIPAIS TONELADA DE FIBRA POR HECTARE COM PALHA (TFHcp) E TONELADA DE SACAROSE POR HECTARE SEM PALHA (TPHsp) EM CANA-PLANTA E PRIMEIRA SOQUEIRA NA SEGUNDA FASE DE TEST.39
- TABELA 4.1 - DESCRIÇÃO DOS CARACTERES AVALIADOS E SUAS RESPECTIVAS NOTAS EM CANA ENERGIA. FONTE: RIDESA..... 51
- TABELA 4.2 - NOTAS ATRIBUÍDAS PARA PRESENÇA DE FERRUGEM MARROM, FERRUGEM ALARANJADA E INCIDÊNCIA DE CARVÃO EM CANA ENERGIA. .... 51
- TABELA 4.3 - COEFICIENTES ESTIMADOS PELO MODELO DE REGRESSÃO LOGÍSTICO PARA VARIÁVEIS QUALITATIVAS (MODELO GERAL QUALITATIVO), NA SELEÇÃO MASSAL REALIZADA NA FASE T2 EM CANA ENERGIA..... 53
- TABELA 4.4 - COEFICIENTES ESTIMADOS PELO MODELO DE REGRESSÃO LOGÍSTICO PARA VARIÁVEIS QUANTITATIVAS (MODELO GERAL

QUANTITATIVO), NA SELEÇÃO MASSAL REALIZADA NA FASE T2  
EM CANA ENERGIA..... 54

TABELA 4.5 - PARÂMETROS ESTIMADOS PARA OS MODELOS REDUZIDOS  
QUALITATIVO E QUANTITATIVO REFERENTE A SELEÇÃO  
MASSAL REALIZADA NA FASE  
T2.....54

TABELA 4.6 - CLASSIFICAÇÃO DO NÚMERO DE GENÓTIPOS SELECIONADOS E  
DESCARTADOS COM A UTILIZAÇÃO DO MODELO DE  
REGRESSÃO LOGÍSTICO PARA VARIÁVEIS QUALITATIVAS E  
QUANTITATIVAS..... 56

TABELA 4.7 - MÉDIA DOS GENÓTIPOS SELECIONADOS (1) E NÃO  
SELECIONADOS (0) PELOS MODELOS QUALITATIVO E  
QUANTITATIVO E PELA SELEÇÃO MASSAL REALIZADA NA FASE  
T2 EM CANA ENERGIA..... 58

## LISTA DE FIGURAS

- FIGURA 3.1 - DIAGRAMAS CAUSAIS ILUSTRANDO O EFEITO DIRETO E INDIRETO DOS CARACTERES DE PRODUÇÃO SOBRE O CARÁCTER PRINCIPAL. A: TONELADA DE COLMOS POR HECTARE COM PALHA E SEM PALHA ( $TCH_{cp/sp}$ ), B: TONELADA DE FIBRA POR HECTARE COM PALHA ( $TFH_{cp}$ ), E C: TONELADA DE SACAROSE POR HECTARE SEM PALHA ( $TPH_{sp}$ ). AS SETAS EM AMBAS AS DIREÇÕES INDICAM ASSOCIAÇÃO MÚTUA ENTRE OS CARACTERES, DETERMINADA PELO COEFICIENTE DE CORRELAÇÃO ( $r_{xy}$ ) E AS SETAS EM ATUAÇÃO DIRETA PARA UM LADO REPRESENTAM O EFEITO DIRETO, DETERMINADO PELO COEFICIENTE DE TRILHA ( $p_{x,y}$ ); AC = ALTURA DO COLMO, DC = DIÂMETRO DO COLMO, NC= NÚMERO DE COLMOS, PMC = PESO MÉDIO DO COLMO, FIB = TEOR DE FIBRA, PC = TEOR DE SACAROSE APARENTE..... 34
- FIGURA 3.2 - MÉDIAS PARA TONELADA DE COLMOS POR HECTARE, TONELADA DE FIBRA POR HECTARE E TONELADA DE SACAROSE POR HECTARE NOS CICLOS DE CANA-PLANTA (A, C, E) E PRIMEIRA SOQUEIRA (B, D, F) EM FUNÇÃO DA SELEÇÃO DIRETA OU INDIRETA E DO NÚMERO DE CLONES SELECIONADOS..... 42
- FIGURA 4.1 - NÚMERO DE CLONES SELECIONADOS PELO MÉTODO MASSAL (A), PELO MODELO QUALITATIVO (B) E QUANTITATIVO (C)..... 57

## LISTA DE ABREVIATURAS

AC - Altura do colmo;

CARV - Incidência de Carvão;

DC - Diâmetro do colmo;

FIB - Teor de fibra;

NC - Número de colmos;

PMCcp - Peso médio do colmo com palha;

PMCsp - Peso médio do colmo sem palha;

PC - Teor de sacarose aparente;

TCHcp - Tonelada de colmos por hectare com palha;

TCHsp - Tonelada de colmos por hectare sem palha;

ACn - Nota para altura do colmo;

CARVn - Nota para incidência de carvão;

CHO - Chochamento;

PB - Presença de brotação lateral;

DCn - Nota para diâmetro do colmo;

DES - Despalha;

HAB - Hábito de crescimento;

FER - Ferrugem marrom;

FAL - Ferrugem alaranjada;

FLO- Florescimento;

PMC - Peso médio de colmos;

SG - Saliência da gema.

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO GERAL</b> .....	<b>12</b>
<b>2</b>	<b>REVISÃO DE LITERATURA</b> .....	<b>15</b>
2.1	SELEÇÃO CLONAL EM CANA-DE-AÇÚCAR.....	15
2.1.1	A cana-de-açúcar e a produção de biomassa .....	16
2.2	CORRELAÇÕES / ASSOCIAÇÕES .....	18
2.2.1	Correlação linear de Pearson .....	18
2.2.2	Regressão linear .....	19
2.3	ANÁLISE DE TRILHA.....	20
2.3.1	Estimativa dos efeitos diretos e indiretos.....	21
2.3.2	Multicolinearidade.....	22
2.4	REGRESSÃO LOGÍSTICA.....	23
2.5	REFERÊNCIAS .....	25
<b>3</b>	<b>CAPÍTULO I - ASSOCIAÇÕES ENTRE CARACTERES AGRONÔMICOS EM CLONES DE CANA ENERGIA</b> .....	<b>29</b>
	RESUMO.....	29
	ABSTRACT .....	30
3.1	INTRODUÇÃO .....	31
3.2	MATERIAL E MÉTODOS .....	32
3.3	RESULTADOS E DISCUSSÃO.....	35
3.4	CONCLUSÕES .....	43
3.5	REFERÊNCIAS .....	43
<b>4</b>	<b>CAPÍTULO II - SELEÇÃO DE CLONES DE CANA ENERGIA VIA MODELO LOGÍSTICO</b> .....	<b>47</b>
	RESUMO.....	47
	ABSTRACT .....	48
4.1	INTRODUÇÃO .....	49
4.2	MATERIAL E MÉTODOS .....	50
4.3	RESULTADOS E DISCUSSÃO.....	53
4.4	CONCLUSÕES .....	59
4.5	REFERÊNCIAS .....	59
<b>5</b>	<b>CONSIDERAÇÕES FINAIS</b> .....	<b>62</b>
<b>6</b>	<b>REFERÊNCIAS GERAIS</b> .....	<b>63</b>
	APÊNDICE 1- Comandos análise de trilha.....	70

APÊNDICE 2 - Comandos análise de regressão logística.....	71
---	----

## 1 INTRODUÇÃO GERAL

Os problemas ambientais gerados pelo consumo de combustíveis fósseis têm levado a uma crescente procura por alternativas sustentáveis, visando reduzir ou complementar a utilização do mesmo, com conseqüente redução na emissão de dióxido de carbono. Além disso, os combustíveis fósseis não são finitos e grandes responsáveis pelo aquecimento global.

A energia contida na biomassa vegetal se torna uma das principais fontes alternativas de energia renovável. Na escolha das culturas que servirão como fonte de biomassa, deve-se considerar aquelas que contribuam com maior eficiência, principalmente não competindo com a produção de alimentos, além de oferecer maior ganho energético e com baixo custo.

A cana-de-açúcar é uma planta que se desenvolve bem nos trópicos e subtropicais, possui o processo fotossintético C4, que consiste em maior eficiência na fixação de carbono em condições de temperaturas mais elevadas. A cultura apresenta alta produtividade, sendo uma das opções para a produção de biomassa, por possuir inúmeras usinas distribuídas pelo país, além do domínio em relação às técnicas de cultivo e processamento.

A busca por novas fontes de energia alternativa fez com que alguns programas de melhoramento genético da cana-de-açúcar direcionassem parte das suas pesquisas para desenvolver variedades com maior teor de fibra. As novas variedades são denominadas de cana energia e visam atender a necessidade do mercado seja para a produção de etanol de segunda geração ou para a cogeração de energia elétrica.

A Rede Interuniversitária para o Desenvolvimento do Setor Sucroenergético (RIDESA), também iniciou um programa de hibridação direcionado para a obtenção de clones que apresentem um maior teor de fibra. O programa de hibridação tem envolvido, principalmente, acessos de *Saccharum spontaneum*, *Saccharum robustum*, variedades atuais da RIDESA (RB's) e de outros programas de melhoramento genético.

Os clones obtidos dos cruzamentos podem ser classificados em três ideótipos. Sendo, cana-de-açúcar convencional, que apresenta valores aproximados de 13% para açúcar e 12% para fibra; cana-energia tipo I, com teor de açúcar pouco menor

ou igual ao convencional, todavia com maior teor de fibra (17%); cana-energia tipo II, com baixa produção de açúcar (5%) e elevado teor de fibra (30%). Os novos clones que serão selecionados para o incremento no teor de fibra são considerados mais rústicos, acarretando em vantagens como menor exigência em água, clima, solo e nutrientes, além da maior capacidade de multiplicação.

A seleção dos clones realizada nas primeiras fases ainda é considerada subjetiva, devido à baixa precisão experimental atrelada ao número reduzido de material propagativo para replicação. O desenvolvimento de um clone de cana-de-açúcar que apresente elevada produtividade de biomassa tornou-se objetivo nos programas de melhoramento de cana energia, logo a seleção dos novos clones de cana energia ainda está passando por processo de aperfeiçoamento.

A utilização de estratégias que visam diminuir custos e tempo com avaliações são importantes e dentre elas está à seleção indireta. Para isso, é importante se ter o conhecimento dos caracteres que influenciam os componentes de produção e os componentes de qualidade nas diferentes fases do melhoramento genético e nos diferentes ciclos da cultura, além de conhecer quais caracteres são mais importantes durante o processo de seleção individual.

O conhecimento das correlações existentes entre os caracteres representa grande importância nos programas de melhoramento genético, principalmente quando a seleção de um, apresenta dificuldade e/ou problemas de medição e identificação, ou por apresentar baixa herdabilidade. Porém, a utilização da análise de correlação não permite obter informações a respeito dos efeitos diretos e indiretos de um grupo de caracteres em relação a um determinado caractere considerado de maior importância. Dessa forma, informações mais úteis no melhoramento genético podem ser obtidas por meio da análise de trilha.

O emprego da análise de trilha nos programas de melhoramento da cana-de-açúcar é de grande valia nas etapas iniciais, pois a metodologia visa apontar os caracteres mais adequados para que seja feita uma seleção indireta dos genótipos mais produtivos, uma vez que a quantificação da produção dos clones é trabalhosa devido ao grande número de clones avaliados nas etapas iniciais.

A análise de regressão logística é outra técnica que também pode ser utilizada para auxiliar nos programas de melhoramento, nas primeiras fases de seleção. A escolha de selecionar ou rejeitar um clone é uma resposta binária, logo, uma análise de regressão logística pode ser aplicada para avaliar o efeito ou poder de

discriminação de cada caractere agronômico utilizado como critério de seleção, bem como a predição de genótipos a serem selecionados.

O objetivo desse trabalho foi determinar a importância dos caracteres agronômicos envolvidos no processo de seleção de clones, visando otimizar o processo de seleção em populações destinadas a obtenção de cana energia.

## 2 REVISÃO DE LITERATURA

### 2.1 SELEÇÃO CLONAL EM CANA-DE-AÇÚCAR

A espécie *Saccharum officinarum* teve sua origem a partir de acessos de *Saccharum robustum*, que apresentavam alto teor de sacarose, baixo teor de fibra e colmos mais grossos. Após passar por processos de “domesticação” em Nova Guiné, selecionando caracteres desejáveis a mesma passou a ser utilizada pelos melhoristas (BORÉM, 2005).

O melhoramento empírico da cana-de-açúcar resultou nas primeiras mudanças gênicas dirigidas, ajudando no processo de evolução das espécies hoje cultivadas. A partir das hibridações realizadas em Java e Barbados em 1888 com cana-de-açúcar foi descoberto que a mesma era capaz de produzir sementes viáveis (CESNIK e MIOCQUE, 2004). Atualmente, os cruzamentos são planejados preferencialmente entre genótipos divergentes, com o intuito de aproveitar o efeito causado pela heterose.

A seleção massal tem sido o método empregado nos programas de melhoramento de cana-de-açúcar. Este método consiste na seleção visual individual de clones, onde as plantas são selecionadas com base em valores fenotípicos (OLIVEIRA et al., 2011), sendo realizada por profissionais com experiência no programa de melhoramento de cana-de-açúcar. Os programas de melhoramento genético realizam inúmeros cruzamentos com o intuito de obter uma combinação que resulte em um clone geneticamente superior, para poder ser lançado como nova variedade (DINARCO-MIRANDA et al., 2008).

A escolha dos genitores e o planejamento dos cruzamentos são importantes etapas para atingir os objetivos de um programa de melhoramento. Na seleção dos genitores, um conjunto de caracteres agronômicos ou o uso de dados de genealogia podem ser considerados (CRESTE et al., 2008). Os cruzamentos entre genitores superiores, muitos dos quais são cultivares clonais já em uso comercial, seguidos por seleção individual, são os procedimentos clássicos adotados no melhoramento de plantas de espécies de propagação assexuada (RESENDE e BARBOSA, 2006).

Um programa de melhoramento genético de cana-de-açúcar envolve cinco fases: três fases de seleção (T1 onde as plântulas obtidas a partir dos cruzamentos são transplantadas a campo, T2 nessa fase ocorre a primeira propagação assexuada

e o material passa a receber o nome de clone, T3 onde ocorre a avaliação dos clones selecionados em T2), uma fase experimental (FE) e uma fase de multiplicação (FM) (BARBOSA e SILVEIRA, 2012). A primeira fase de seleção (T1) é considerada uma das mais importantes de um programa de melhoramento (OLIVEIRA et al., 2011), uma vez que nesta fase são selecionados os primeiros materiais, nas outras fases são usados os materiais selecionados em T1. Por isso é importante ter uma seleção eficiente, para poder atingir os objetivos.

Na escolha dos clones, vários critérios são levados em consideração, como um conjunto de caracteres, para se ter uma avaliação mais precisa do indivíduo. A definição do critério de seleção depende dos parâmetros genéticos e fenotípicos (herdabilidades, repetibilidades e correlações genéticas e fenotípicas) que estão associados aos caracteres (RESENDE, 2002).

A seleção realizada principalmente nas duas primeiras etapas (T1 e T2), apresenta menor eficiência, comparada com as demais etapas. Isso acontece em razão das baixas herdabilidades, no sentido amplo, para a maioria dos caracteres (LANDELL e BRESSIANI, 2005).

### 2.1.1 A cana-de-açúcar e a produção de biomassa

Os programas de melhoramento genético de cana-de-açúcar sempre direcionaram as pesquisas para obtenção de genótipos com alto teor de sacarose (ZENI NETO et al., 2013; DALBIANCO et al., 2012). Nos últimos anos, trabalhos envolvendo a utilização de biomassa oriunda de cana-de-açúcar como fonte alternativa de energia tem aumentado, com o intuito de tentar diminuir a utilização de combustíveis fósseis (WACLAWOVSKY et al., 2010).

A biomassa vegetal passou a ser uma fonte alternativa na redução da utilização de combustíveis fósseis. Na escolha da fonte de biomassa alguns atributos devem ser considerados, como, a planta apresentar alto rendimento, crescimento rápido, com parede celular de fácil quebra e alta eficiência no crescimento e com baixo requerimento de energia na colheita (WACLAWOVSKY et al., 2010).

As culturas energéticas não devem exigir o uso extensivo de terras agrícolas e devem ter baixo custo de produção de energia a partir da biomassa. Nesse sentido, a cana-de-açúcar torna-se uma das culturas promissoras pois possui a maioria dos

atributos acima, podendo ser utilizada como fonte alternativa para redução da utilização de combustíveis fósseis. (WACLAWOVSKY et al., 2010).

A necessidade de uma quantidade maior de biomassa para produzir etanol de segunda geração e energia elétrica possibilitou o surgimento de novas pesquisas voltadas para o desenvolvimento de variedades de cana-de-açúcar com maior teor de fibra, denominadas de cana energia (LOUREIRO et al., 2011; SILVEIRA et al., 2015; RAMOS et al., 2017).

Os cruzamentos envolvendo, principalmente, as espécies *Saccharum robustum*, *Saccharum spontaneum* e híbridos comerciais de cana-de-açúcar, resultando em variedades mais robustas, com maior teor de fibra e potencial produtivo, são denominadas cana energia (MATSUOKA et al., 2010), podendo ser utilizada na fabricação de etanol de segunda geração e/ou cogeração de energia elétrica.

O etanol de segunda geração oriundo, de cana-de-açúcar ainda não foi usado comercialmente. Mas o germoplasma dos programas de melhoramento está sendo explorado, não apenas para o rendimento de sacarose, mas também para o rendimento de biomassa, antecipando tecnologias que podem permitir uma produção eficiente de energia a partir de resíduos celulósicos (MATSUOKA et al., 2010).

A produção de etanol de segunda geração em larga escala é um objetivo altamente desejado, porque poderia ampliar significativamente a escolha de matérias-primas. Seria então possível utilizar biomassa de culturas, que não fossem alimentos para animais ou alimentos para humanos, para a geração de biocombustíveis (TOLLEFSON, 2008).

Os colmos da cana-de-açúcar são compostos basicamente por fibra substância de reserva energética e sacarose. O colmo de uma variedade produz de 10 a 14% de fibra e de 12 a 17% de sacarose. O peso de cada colmo adulto de cana-de-açúcar constitui-se de aproximadamente 15% de palha, folhas e ponteiros, sendo que 40% desse resíduo não é utilizado na co-geração de energia (CARVALHO-NETTO et al., 2014).

A produção anual de cana-de-açúcar por hectare gera em média 39 t/ha<sup>-1</sup> de resíduo, sendo comparada favoravelmente com outras culturas de alto rendimento como a *Miscanthus* (29,6 t/ha<sup>-1</sup>) e milho (total de grãos mais forragem, 17,6 t/ha<sup>-1</sup>) (HEATON et al., 2008; WACLAWOVSKY et al., 2010). A extração do caldo contido

nos colmos durante o processo de moagem gera em média 240 kg/t de bagaço, resíduos do processo (ENSINAS et al., 2007).

## 2.2 CORRELAÇÕES / ASSOCIAÇÕES

As causas de correlação entre características podem ser genéticas ou ambientais. As correlações genéticas estão relacionadas a fatores genéticos da população analisada (CRUZ et al., 2014), já a correlação fenotípica se refere a associação entre duas variáveis que podem ser observadas diretamente (FALCONER, 1987).

A análise de correlação pode ser utilizada para medir a existência e intensidade de interação entre caracteres, ou seja, o grau de associação linear entre duas variáveis quaisquer (FERREIRA, 2007). Essa análise tem sido utilizada em estudos de melhoramento genético, pois permite a realização de seleção indireta de uma característica de difícil mensuração por meio de outras características que estejam correlacionadas, dessa forma levando ao progresso do caráter desejado (CRUZ, 2006).

O coeficiente de correlação estabelece a existência ou ausência de relação entre duas variáveis, sem necessitar do ajuste de uma função matemática. O coeficiente de correlação linear mostra que a mudança de uma variável implica na mudança da outra variável e pode ser positivo quando ambas as variáveis aumentam ou diminuem e negativo quando uma variável aumenta e a outra diminui (SIEGEL e CASTELLAN, 2006).

### 2.2.1 Correlação linear de Pearson

O coeficiente de correlação linear de Pearson está compreendido no intervalo de  $[-1,1]$ , valores próximos de 1 indicam correlação forte entre as variáveis, enquanto valores próximos de 0 indicam baixa correlação. O sinal indica se a correlação é positiva quando ocorre relação direta entre as variáveis ou negativo quando há uma associação inversa entre elas (CASELLA e BERGER, 2001).

O coeficiente de correlação entre duas variáveis  $x$  e  $y$  é fornecido pela expressão:

$$r_{xy} = \frac{Cov(x, y)}{\sqrt{V(x) \times V(y)}} = \frac{\sigma_{x,y}}{\sigma_x \times \sigma_y}$$

Onde:  $\sigma_{x,y}$  é a covariância entre as variáveis aleatórias  $x$  e  $y$ ;  $\sigma_x$  é o desvio padrão da variável aleatória  $x$ ; e o  $\sigma_y$  é o desvio padrão da variável aleatória  $y$ .

O teste de significância para a correlação entre as variáveis  $x$  e  $y$  é analisado pela seguinte estatística de significância (COSTA NETO, 2009):

$$t = \frac{r_{xy}\sqrt{n-2}}{\sqrt{1-r_{xy}^2}}$$

Em que:  $r_{xy}$  é a correlação linear entre  $x$  e  $y$ , e  $n$  é o número de observações na amostra. A hipótese de nulidade  $H_0(H_0: \rho = 0$ , a correlação é estatisticamente igual a zero) não é rejeitada se  $|t| < t_{(\frac{\alpha}{2}; n-2)}$  (COSTA NETO, 2009).

### 2.2.2 Regressão linear

A análise de regressão estuda a relação de uma variável dependente e outra independente através de um modelo matemático. A análise de correlação é estreitamente relacionada com a análise de regressão. Essas técnicas analisam como duas ou mais variáveis podem estar relacionadas. A correlação fornece o número que resume a magnitude e a análise de regressão fornece um modelo matemático que descreve essa relação. O modelo obtido na análise de regressão pode ser utilizado para estimar ou prever valores de uma variável, quando se conhece ou supõe os valores de outra variável (COSTA NETO, 2009).

O modelo de regressão linear simples se define a partir de uma relação linear entre a variável dependente e uma variável independente. Se existirem várias variáveis independentes, o modelo passa a denominar-se modelo de regressão linear múltipla (CHARNET et al., 2008). A regressão linear múltipla envolve uma única variável dependente e duas ou mais variáveis independentes. A análise tem por objetivo encontrar um modelo, que possa ser utilizado para prever os valores de  $Y$  dado valores das diversas variáveis independentes.

Os parâmetros ( $\beta_0, \beta_1, \beta_2, \beta...$ ) da regressão podem ser estimados pelo método dos mínimos quadrados e também pode se utilizar o método da máxima verossimilhança. Após a estimação dos parâmetros, há a necessidade de testar a significância da regressão e realizar um teste individual para cada coeficiente, a fim de validar o modelo. Para utilizar o método de máxima verossimilhança e realizar as inferências estatísticas, deve-se supor que exista normalidade dos resíduos (CHARNET et al., 2008; MONTGOMERY et al., 2012).

No ajuste de modelos de regressão linear múltipla deve-se detectar a presença de variáveis independentes multicolineares. A violação desta orientação implica risco de inferências espúrias (MONTGOMERY et al., 2012). A qualidade do modelo ajustado pode ser feita utilizando diversas medidas como o coeficiente de determinação ( $R^2$ ), soma de quadrado de resíduos, etc. (CHARNET et al., 2008; HAIR et al., 2009; MONTGOMERY et al., 2012).

### 2.3. ANÁLISE DE TRILHA

A análise de correlação não nos permite identificar o real efeito que cada caractere explicativo (independente) apresenta sobre o caractere resposta (dependente). Por isso, a análise de trilha descrita por Wright (1921) é utilizada para proporcionar uma melhor compreensão das associações entre os caracteres, a partir do desdobramento das correlações simples. Essa metodologia permite quantificar a grandeza e a direção das associações entre caracteres complexos, mostrando a importância dos efeitos diretos e indiretos sobre o caractere de interesse (TYAGI e LAL, 2007; CRUZ et al., 2014).

Os coeficientes de trilha obtidos quantificam os efeitos diretos e indiretos. Os mesmos são obtidos por meio de equações de regressão, sendo as variáveis previamente padronizadas. Essas estimativas são obtidas a partir do método de mínimos quadrados. Quando se considera um único modelo casual, a análise de trilha trata-se de uma análise de regressão múltipla padronizada (CRUZ et al., 2014).

A análise consiste na construção de um diagrama de trilha que irá ser útil para exibir graficamente o padrão de hipótese das relações de causa e efeito, seguido da decomposição das correlações em coeficientes de um caractere sobre o outro.

No melhoramento genético, a análise de trilha, pode ser utilizada para determinar a importância de caracteres primários e secundários da cultura sobre a

produtividade, possibilitando o uso da seleção indireta (CRUZ et al., 2014). Por isso, a compreensão das associações entre os caracteres é imprescindível ao melhoramento genético, pois auxilia na escolha da melhor estratégia durante o processo de seleção.

### 2.3.1 Estimativa dos efeitos diretos e indiretos

Considerando uma variável  $Y$  como a resposta e as variáveis  $X_1, X_2, X_3$  explicativas, relacionam-se por meio de um único modelo de regressão múltipla:

$$Y = b_{yx_1}(X_1) + b_{yx_2}(X_2) + b_{yx_3}(X_3) + \varepsilon$$

Onde:

$Y$  = variável dependente

$X_1, X_2, X_3$  = variáveis independentes

$b_{yx_1}, b_{yx_2}, b_{yx_3}$  = parâmetros estimados

A padronização das variáveis é feita pela subtração da média de cada variável seguido da divisão de ambos os membros pelo desvio padrão da variável básica e posterior multiplicação e divisão de cada termo do segundo membro pelo respectivo desvio padrão da variável associada a esse termo (LI, 1975):

$$\frac{Y-\bar{Y}}{\sigma_Y} = \frac{b_{YX_1}(X_1-\bar{X}_1)}{\sigma_Y} \cdot \frac{\sigma_{X_1}}{\sigma_{X_1}} + \frac{b_{YX_2}(X_2-\bar{X}_2)}{\sigma_Y} \cdot \frac{\sigma_{X_2}}{\sigma_{X_2}} + \frac{b_{YX_3}(X_3-\bar{X}_3)}{\sigma_Y} \cdot \frac{\sigma_{X_3}}{\sigma_{X_3}} + \frac{(\varepsilon-\bar{\varepsilon})}{\sigma_Y} \cdot \frac{\sigma_\varepsilon}{\sigma_\varepsilon}$$

As estimativas dos efeitos diretos e indiretos por meio da análise de trilha é realizado considerando-se o seguinte modelo:

$$y = p_{yx_1}x_1 + p_{yx_2}x_2 + p_{yx_3}x_3 + p_\varepsilon u$$

Sendo:

$y = \frac{Y-\bar{Y}}{\sigma_Y}$  variável básica padronizada

$x_i = \frac{X_i-\bar{X}_i}{\sigma_{X_i}}$  variável explicativa padronizada

$u = \frac{\varepsilon-\bar{\varepsilon}}{\sigma_\varepsilon}$ ;  $p_\varepsilon = \frac{\sigma_\varepsilon}{\sigma_Y}$  coeficiente da variável residual na análise de trilha e

$p_{yx_i} = \frac{b_{YX_i} \sigma_{X_i}}{\sigma_Y}$  coeficiente da variável explicativa na análise de trilha

As estimativas dos coeficientes de trilha são obtidas pela resolução do sistema linear de equações  $\bar{Y} = \bar{X}\bar{P}$ , onde  $\bar{X}$  é a matriz não singular das correlações entre as variáveis explicativas;  $\bar{P}$  um vetor coluna contendo as estimativas dos coeficientes e  $\bar{Y}$  é o vetor coluna das correlações entre a variável principal e cada variável explicativa do modelo.

$$\bar{Y} = \begin{bmatrix} r_{YX_1} \\ r_{YX_2} \\ r_{YX_3} \end{bmatrix}, \bar{X} = \begin{bmatrix} 1 & r_{12} & r_{13} \\ r_{21} & 1 & r_{23} \\ r_{31} & r_{32} & 1 \end{bmatrix} \text{ e } \bar{P} = \begin{bmatrix} p_{xy_1} \\ p_{xy_2} \\ p_{xy_3} \end{bmatrix}$$

### 2.3.2 Multicolinearidade

O termo multicolinearidade significa a existência de uma correlação linear exata entre duas ou mais variáveis e foi proposto em 1934 por Ragnar Frich (KUTNER et al., 2005). É difícil de ocorrer uma correlação exata entre duas variáveis, logo, o termo multicolinearidade é utilizado em casos de correlação alta entre os caracteres.

Na análise de trilha o que interessa é a multicolinearidade existente entre os caracteres independentes, pois as altas correlações dificultam a estimação do caractere explicativo, por estarem associados e contribuírem de forma coletiva para explicar o caráter dependente. A presença de multicolinearidade faz com que a variância associada aos estimadores de coeficientes atinja valores altos, tornando as estimativas pouco confiáveis (CRUZ et al., 2014).

A colinearidade existente entre os caracteres independentes é nociva no ajuste de modelos de regressão linear múltipla, logo a existência de multicolinearidade nas variáveis explicativas pode vir comprometer os resultados da análise de trilha, em razão da mesma estar inserida nos modelos de regressão linear múltipla.

## 2.4 REGRESSÃO LOGÍSTICA

A maioria das análises baseadas em modelos lineares assumem dados normalmente distribuídos da variável resposta e dos termos de erro dos modelos ajustados. A natureza binária da variável dependente tem propriedades que violam algumas suposições da regressão múltipla. O termo de erro de uma variável discreta segue distribuição binomial ao invés da normal e, além disso, a variância de uma variável dicotômica não é constante, criando casos de heterocedasticidade (HAIR et al., 2009).

A análise de regressão logística foi proposta no final de 1960 e início da década de 1970 (CABRERA, 1994) sendo inicialmente utilizada na pesquisa educacional, especialmente no ensino superior (AUSTIN, YAFFEE e HINKLE, 1992; CABRERA, 1994; PENG e SO, 2002; PENG et al., 2002).

Os problemas relacionados a pesquisa educacional demandam a análise e previsão de um resultado dicotômico: se um estudante terá sucesso na faculdade, se uma criança deve ser classificada como com deficiência de aprendizagem, se um adolescente é propenso a se envolver em comportamentos de risco, e assim por diante. Tradicionalmente, essas perguntas de pesquisa foram abordadas por regressão de mínimos quadrados ordinários (PENG et al., 2013).

A técnica de mínimos quadrados foi posteriormente discutida como ineficiente para predição de resultados dicotômicos, em razão das suas suposições estatísticas estritas, exigindo linearidade e normalidade dos dados (COX e SNELL, 1989, LEI e KOEHLI, 2000, TABACHNICK e FIDELL, 2001).

Os modelos de regressão constituem uma das ferramentas estatísticas mais importantes na análise de dados quando se pretende modelar relações entre variáveis. O objetivo desses modelos é explorar a relação entre uma ou mais variáveis explicativas e uma variável resposta (AFIFI et al., 2012). Um dos casos particulares dos modelos lineares generalizados são os modelos onde a variável resposta apresenta apenas duas categorias, assumindo valores 0 ou 1, sendo o modelo de regressão logística o mais popular (AGRESTI., 2007).

Os modelos de regressão logística são parte dos modelos lineares generalizados e são utilizados para prever a probabilidade de ocorrência de eventos binários, ajustando os dados das variáveis preditoras a uma curva logística (AGRESTI, 2007; PRESS, 1978). Na área de melhoramento de plantas, a variável

resposta pode ser a decisão de selecionar ou rejeitar um genótipo. A função linear das variáveis preditoras constitui uma regressão linear múltipla. A equação de regressão logística a partir da qual a probabilidade da variável  $Y$  é prevista é dada por:

$$\pi(x) = \frac{e^{(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n)}}{1 + e^{(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n)}}$$

Onde  $\pi(x)$  é a probabilidade de selecionar uma planta,  $\beta_0$  é a intercepção da função preditora,  $\beta_1$  é o coeficiente da variável 1,  $\beta_2$  é o coeficiente da variável 2,  $\beta_n$  é o coeficiente da  $n$ ésima variável.

A regressão logística permite modelar, a partir de um conjunto de observações, a relação “logística” entre uma variável resposta dicotômica e uma série de variáveis explicativas numéricas, sejam elas contínuas discretas e/ou categóricas. O modelo apresenta grande potencial no processo de seleção individual dentro das melhores famílias em cana-de-açúcar, auxiliando na identificação dos melhores genótipos (BRASILEIRO et al., 2015).

## 2.5 REFERÊNCIAS

- AFIFI, A.; MAY, S.; CLARK, V.A. **Practical Multivariate Analysis**. 5ª ed, 2012.
- AGRESTI, A. **An introduction to categorical data analysis**. 2ª ed. Hoboken: John Wiley & Sons, 2007.
- AUSTIN, J.T.; YAFFEE, R.A.; HINKLE, D.E. Logistic regression for research in higher education. **Higher Education: Handbook of Theory and Research**, v.8, p.379-410, 1992.
- BARBOSA, M.H.P.; SILVEIRA, L.C.I. Breeding and cultivar recommendations. In Santos F, Borém A and Caldas C (eds.) **Sugarcane: bioenergy, sugar and ethanol: technology and prospects**. MAPA/ACS: UFV/DEA, Brasília, p. 301-318, 2012.
- BRASILEIRO, B.P.; MARINHO, C.D.; COSTA, P.M.A.; CRUZ, C.D.; PETERNELLI, L.A.; BARBOSA, C.H.P. Selection in sugarcane families with artificial neural networks. **Crop Breeding and Applied Biotechnology** v.15, p.72-78, 2015.
- BORÉM, A. **Melhoramento de espécies cultivadas**. Viçosa: UFV 2 ed, 969p, 2005.
- CABRERA, A.F. Logistic regression analysis in higher education: An applied perspective. **Higher Education: Handbook of Theory and Research, Vol. 10**, 225-256, 1994.
- CARVALHO-NETTO, O.V.; BRESSIANI, J.A.; SORIANO, H.L.; FIORI, C.S.; SANTOS, J.M.; BARBOSA, G.V.S.; XAVIER, M.A.; LANDELL, M.G.A.; PEREIRA, G.A.G. The potential of the energy cane as the main biomass crop for the cellulosic industry. **Chemical and Biological Technologies in Agriculture**, v.1, p.1-8, 2014.
- CASELLA, G.; BERGER, R. **Interferência estatística**. 2<sup>nd</sup> ed. São Paulo: C Learning, 588p, 2001.
- CESNIK, R.; MIOCQUE, J. **Melhoramento da cana-de-açúcar**. Embrapa Informação Tecnológica, 307.p, 2004.
- CHARNET, R. **Análise de modelos de regressão linear**. Campinas: Unicamp, 357p, 2008.
- COSTA NETO, P.L.O. **Estatística**. 2. Ed. São Paulo: E. Blucher, 280p, 2009.
- COX, D.R.; SNELL, E.J. **The analysis of binary data** (2nd ed.). London: Chapman and Hall, 1989.
- CRESTE, S.; ROSA JUNIOR, V.E.; PINTO, L.R.; ALBINO, J.C.; FIGUEIRA, A.V.O. A Biotecnologia como ferramenta para o melhoramento genético. In: DINARDO MIRANDA, L.L.; VASCONCELOS, A.C.M.; LANDELL, M.G.A. **Cana-de-açúcar**. Instituto agrônomo. p.156-176. 2008.

CRUZ, C.D. Programa GENES: **análise multivariada e simulação**. Viçosa: UFV, 175p, 2006.

CRUZ, C.D.; CARNEIRO, P.C.S.; REGAZZI, A.J. **Modelos biométricos aplicados ao melhoramento genético**. 3ª ed. UFV, Vicosa, MG, Brazil, 2014.

DAL-BIANCO, M.; CARNEIRO, M.S.; HOTTA, C.T.; CHAPOLA, R.G.; HOFFMANN, H.P.; GARCIA, A.A.F.; SOUZA, G.M. Sugarcane improvement: how far can we go?. **Elsevier Current Opinion in Biotechnology**, v.23, p.265-270. 2012.

DINARDO-MIRANDA, L.L.; VASCONCELOS, A.C.M.; LANDELL, M.G.A. **Cana-de-açúcar**. Campinas: Instituto Agrônômico, 882.p, 2008.

ENSINAS, A.V.; NEBRA, S.A.; LOZANO, M.A.; SERRA, L.M. Analysis of process steam demand reduction and electricity generation in sugar and ethanol production from sugarcane. **Energy Conversion and Management**, v.48, p.2978-2987, 2007.

FALCONER, D.S. **Introdução a genética quantitativa**. Viçosa, UFV, Impr.Univ, 279p, 1987.

FERREIRA, F.M.; BARROS, W.S.; SILVA, F.L.; BARBOSA, M.H.P.; CRUZ, C.D.; BASTOS, I.T. Relação fenotípica e genotípica entre componentes de produção em cana-de-açúcar. **Bragantia**, v.66, n.4, p.605-610, 2007.

HAIR, J.F.; BLACK, W.C.; BABIN, B.J.; ANDERSON, R.E.; TATHAM, R.L. **Análise Multivariada de Dados**. Tradução Adonai Schlup Sant'Anna. -6. Ed.- Porto Alegre, 688.p, 2009.

HEATON, E.A.; DOHLEMAN, F.G.; LONG, S.P. Meeting US biofuel goals with less land: the potential of Miscanthus. **Global Chang Biology**, v.14, p.2000-2014, 2008.

LANDELL, M.G.A.; BRESSIANI, J.A. **Melhoramento genético, caracterização e manejo varietal**. In: DINARDO-MIRANDA, L.L.; VASCONCELOS, A.C.M.; LANDELL, M.G.A. **Cana-de-açúcar**. Campinas, SP: Instituto Agrônômico. p.101-155. 2005.

LEI, P.W.; KOEHLI, L.M. **Linear discriminant analysis versus logistic regression: A comparison of classification errors**. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA, 2000.

LI, C.C. Concept of path coefficient and its impact on population genetics. **Biometrics**, v.12, p.190-210, 1975.

LOUREIRO, M.E.; BARBOSA, M.H.P.; LOPES, F.J.F.; SILVÉRIO, F.O. Sugarcane breeding and selection for more efficient biomass conversion in cellulosic ethanol. In: Buckeridge MS, Goldman GH (eds) **Routes to cellulosic ethanol**. Springer, New York, pp 199-239, 2011.

MATSUOKA, S.; BRESSIANI, J.; MACCHERONI, W.; FOUTO, I. Bioenergia de cana. In: SANTOS, F.; BORÉM, A.; CALDAS, C. (Org). **Cana-de-açúcar: Bioenergia, Açúcar e Álcool**. 1ed. Viçosa: UFV, v1. p:547-577, 2010.

MONTGOMERY, D.C.; PECK, E.A.; VINING, G.G. **Introduction to linear regression analyses**. Wiley, p. 504, 2012.

KUTNER, M.H.; NACHTSHEIM, J.C.; NETER, J.; LI, W. **Applied linear statistical models**. 5<sup>th</sup> ed. New York: McGraw- Hill/Irwin, 1396 p, 2005.

OLIVEIRA, R.A.; DAROS, E.; RESENDE, M.D.V.; BESPALHOK-FILHO, J.C.; ZAMBON, J.L.C.; SOUZA, T.R.; LUCIUS, A.S.F. Procedimento Blupis e seleção massal em cana-de-açúcar. **Bragantia**, v. 70, n. 4, p.796-800, 2011.

PENG, C.Y.J.; LEE, K.L.; INGERSOLL, G.M. An introduction to logistic regression analysis and reporting. **The Journal of Educational Research**, v.96, n.1, p. 3-14, 2013.

PENG, C.Y.; SO, T.S.H. Modeling strategies in logistic regression. **Journal of Modern Applied Statistical Methods**, v.14, p.147-156, 2002.

PENG, C.Y.; SO, T.S.; STAGE, F.K.; ST. JOHN, E.P. The use and interpretation of logistic regression in higher education journals: 1988-1999. **Research in Higher Education**, v. 43, p.259-293, 2002.

PRESS, S.J.; WILSON, S. Choosing between logistic regression and discriminant analysis. **Journal of the American Statistical Association**. 73, 699-705, 1978.

RAMOS, R.S.; BRASILEIRO, B.P; SILVEIRA, L.C.I.; KIST, V.; PETERNELLI, L.A.; BARBOSA, M.H.P. Selection of parents, families and clones to obtain energy cane. **Agronomy Journal**, v.109, p.1-7, 2017.

RESENDE, M.D.V.; BARBOSA, M.H.P. Selection via simulated BLUP based on family genotypic effects in sugarcane. **Pesquisa Agropecuária Brasileira**, v.41, p.421-429, 2006.

RESENDE, M.D.V. Genética biométrica e estatística no melhoramento de plantas perenes. **Embrapa Informação Tecnológica**: Brasília, 975p, 2002.

SIEGEL, S.; CASTELLAN JR, N.J. **Estatística não-paramétrica: para ciência do comportamento**. Porto Alegre: Artmed 2<sup>a</sup> Ed, 448p, 2006.

SILVEIRA, L.C.I.; BRASILEIRO, B.P.; KIST, V.; WEBER, H.; DAROS, E.; PETERNELLI, L.A.; BARBOSA, M.H.P. Selection strategy in families of energy cane based on biomass production and quality traits. **Euphytica**, v. 204, p. 443-455, 2015.

TABACHNICK, B.G.; FIDELL, L.S. **Using multivariate statistics** (4<sup>a</sup> ed.). Needham Heights, MA: Allyn & Bacon, 2001.

TOLLEFSON, J. Energy: not your father's biofuels. **Nature**, v.451, p.880-883, 2008.

TYAGI, A.P.; LAL, P. Correlation and path coefficient analysis in sugarcane. **The South Pacific Journal of Natural Science**, v.25, p.1-9, 2007.

ZENI-NETO, H, DAROS, E, BESPALHOK-FILHO, J.C, SCAPIM C.A, VIDIGAL, M.C.G, FILHO, P.S.V. Selection of families and parents of sugarcane (*Saccharum* spp.) through mixed models by joint analysis of two harvests. **Euphytica**, v.193, p.391-408, 2013.

WACLAWOVSKY, A.J.; SATO, P.M.; LEMBKE, C.G.; MOORE, P.H.; SOUZA, G.M. Sugarcane for bioenergy production: an assessment of yield and regulation of sucrose content. **Plant Biotechnology Journal**, v.8, p. 263-276, 2010.

WRIGHT, S. Correlation and causation. **Journal of Agricultural Research**, v.20, n.7, p.557-585, 1921.

### 3 CAPÍTULO I - ASSOCIAÇÕES ENTRE CARACTERES AGRONÔMICOS EM CLONES DE CANA ENERGIA

#### RESUMO

O conhecimento dos caracteres que influenciam os componentes de produção e a qualidade em cana energia é fundamental no processo de seleção de clones dos programas de melhoramento genético. O objetivo desse trabalho foi compreender as associações existentes entre os caracteres componentes da produção e os seus efeitos na tonelada de colmos por hectare com palha e sem palha, na tonelada de fibra por hectare com palha e na tonelada de sacarose por hectare sem palha, em clones de cana energia da segunda fase de teste nos ciclos de cana-planta e primeira soqueira. Foram realizadas avaliações em 220 clones, no ciclo de cana-planta e primeira soqueira. O experimento foi implantado no delineamento em blocos incompletos. Os caracteres mensurados foram os seguintes: diâmetro do colmo, número de colmos, altura do colmo, peso médio do colmo com palha e sem palha, teor de sacarose aparente e o teor de fibra. Foi realizada análise de correlação de Pearson a partir das médias genéticas, estimadas por meio de modelos mistos, seguido pelo desmembramento dos efeitos diretos e indiretos via análise de trilha. O peso médio do colmo com palha, peso médio do colmo sem palha e o número de colmos foram os caracteres que apresentaram maior efeito direto na tonelada de colmos por hectare com palha, na tonelada de sacarose por hectare sem palha e na tonelada de fibra por hectare com palha nos ciclos de cana-planta e primeira soqueira. O peso médio do colmo apresentou correlação negativa com o número de colmos, indicando que a seleção indireta deve ser realizada utilizando apenas um dos dois caracteres. A seleção de clones para teor de fibra e teor de sacarose pode ser realizada via número de colmos, enquanto que, para a tonelada de colmos por hectare a seleção indireta não é vantajosa.

Palavras-chave: *Saccharum* spp., biomassa, análise de trilha, melhoramento de plantas.

## ASSOCIATIONS BETWEEN AGRONOMIC TRAITS IN CLONES OF ENERGY CANE

### ABSTRACT

The knowledge of the traits that influence the components of production and quality in cane energy makes it possible to aid in the selection of clones in breeding programs. The objective of this work was to understand the associations between the component characters of production and their effects on the ton of ton of cane per hectare with straw and without straw, ton of fiber per hectare with straw and ton of sucrose per hectare without straw for the second test phase in the plant cane and first ratoon cycles. The experiment with 220 clones was implanted in an incomplete block design, the evaluations were carried out in the planta cane and first ratoon cycles, the following traits: stalk diameter, stalk number, stalk height, stalk weight with straw and stalk weight without straw, apparent sucrose content and fiber content. Pearson correlation analysis was performed from the genetic means, estimated by means of mixed models, followed by the dismemberment of the direct and indirect effects via track analysis. The average straw weight, straw weight without straw and stalk number were the characters that had the greatest direct effect under a ton of stalks per hectare with straw, a ton of sucrose per hectare without straw and a ton of fiber per hectare straw in the plant cane and first ratoon cycles. The mean weight of the stalk correlates negatively with the number of stalks, indicating that the indirect selection should be performed using only one of the two characters. The selection of clones for fiber content and sucrose content can be performed via number of stalks, per ton of stalks per hectare indirect selection is not advantageous.

Keywords: *Saccharum* spp., biomass, path analysis, plant breeding.

### 3.1 INTRODUÇÃO

A procura por novas fontes de energias limpas e renováveis tem sido constante no mundo. Nos últimos anos aumentaram-se os estudos relacionados ao uso da biomassa para a geração de energia (SILVEIRA et al., 2015a). A necessidade de maior quantidade de biomassa para a produção de etanol de segunda geração e de energia elétrica alavancou o surgimento de novas pesquisas para o desenvolvimento de variedades de cana-de-açúcar com alto teor de fibra, denominadas de cana energia (LOUREIRO et al., 2011; RAMOS et al., 2017).

O ciclo produtivo da cultura e as diferentes fases de seleção envolvidas no processo de melhoramento genético da cana-de-açúcar exigem de 10 a 15 anos para o lançamento de novas variedades, gerando demanda por mão-de-obra, insumos e área experimental (BARBOSA e SILVEIRA, 2012), fazendo com que os programas de melhoramento genético utilizem estratégias para a seleção de clones, para tornar o processo mais eficiente.

O conhecimento das correlações entre os caracteres envolvidos com a produção de biomassa e fibra e os seus efeitos sobre a produtividade podem auxiliar na escolha dos genótipos. A seleção de caracteres que apresentem influência direta sobre caracteres de baixa herdabilidade e/ou dificuldades de medição, possibilita a seleção indireta, com base em um caractere de fácil mensuração e maior herdabilidade, permitindo ao melhorista obter progressos mais rápido em relação ao uso da seleção direta (MARCHEZAN et al., 2005).

A utilização da correlação de Pearson permite estimar o grau de associações entre dois caracteres. Entretanto, a mesma não representa uma medida de causa e efeito e uma interpretação direta das magnitudes pode resultar em equívocos, pois uma alta correlação entre duas variáveis pode ser resultado do efeito de uma ou um conjunto de caracteres (FERREIRA et al., 2007).

Os estudos envolvendo a aplicação da análise de trilha permitem a partição dos coeficientes de correlação linear em efeitos diretos e indiretos de vários caracteres considerados como explicativo, em direção a um caractere dependente (OLIVOTO et al., 2017). A técnica tem auxiliado na seleção indireta no melhoramento genético das culturas (BELLO et al., 2010; NARDINO et al., 2016). No melhoramento de cana-de-açúcar a técnica tem sido altamente vantajosa nas fases iniciais, pois a

metodologia visa indicar os caracteres mais adequados para seleção de clones produtivos (SILVA et al., 2009).

A análise de trilha tem sido utilizada em diversos trabalhos de cana-de-açúcar (CHAUDHARY e JOSHI, 2005; FERREIRA et al., 2007; SILVA et al., 2009; ESPÓSITO et al., 2012; BRASILEIRO, PETERNELLI e BARBOSA, 2013; BAFFA et al., 2014; TENA et al., 2016). Em cana energia há poucas informações, apenas o trabalho de Silveira, (2015b) traz informações para a primeira fase de teste (T1) e em estudos de família. Na segunda fase de teste (T2) não se tem informações de quais caracteres têm maior efeito sobre a produtividade, nos ciclos de cana-planta e primeira soqueira.

Portanto, o objetivo desse trabalho é compreender as associações existentes entre os caracteres componentes da produção e seus efeitos na tonelada de cana por hectare com palha (TCHcp) e sem palha (TCHsp), na tonelada de fibra por hectare com palha (TFHcp) e na tonelada de sacarose por hectare sem palha (TPHsp), em clones de cana energia da segunda fase de teste nos ciclos de cana-planta e cana-soca.

### 3.2 MATERIAL E MÉTODOS

O experimento foi realizado na Estação Experimental de Paranaíba, pertencente a Universidade Federal do Paraná, com latitude de 23°05'S; longitude de 52°27'W e altitude de 503 m. O clima da região é classificado como Cfa segundo Köppen com precipitação anual de 1.200 a 1.400 mm e temperatura média anual entre 22 e 23°C (APARECIDO et al., 2016) e solo Latossolo Vermelho distrófico (SANTOS et al., 2013).

Foram avaliados 220 clones, o experimento foi instalado em junho de 2014 no delineamento em blocos incompletos e cada bloco foi composto por 22 clones. As parcelas foram constituídas de duas linhas de 5 m, espaçadas em 1,4 m e com plantio de 18 gemas por metro. Foram realizadas avaliações no final dos ciclos de cana-planta e primeira soqueira, em junho de 2015 e julho de 2016, respectivamente.

Os caracteres avaliados foram: diâmetro do colmo (DC), com a medição em centímetros no terço médio de 10 colmos por parcela, que foram coletados aleatoriamente; número de colmos (NC), a partir da contagem de colmos presentes na parcela; altura do colmo (AC), com a medição de 10 colmos por parcela, a partir

do nível do solo até o ápice da planta; peso médio do colmo (PMC), com a pesagem de 10 colmos por parcela, com a presença de palha (PMCcp), e posteriormente sem palha (PMCsp).

Os componentes de qualidade, teor de sacarose aparente (PC) e teor de fibra (FIB) foram obtidos a partir da análise tecnológica, seguindo as normas técnicas de determinação de qualidade da matéria prima, conforme recomendado por Fernandes (2011).

A partir dos dados de NC, PMC com palha (PMCcp) e sem palha (PMCsp), PC e FIB foram estimados a tonelada de colmos por hectare (TCH) com palha (TCHcp) e sem palha (TCHsp):  $TCH_{cp/sp} = \frac{(NC \times PMC_{cp/sp} \times 10)}{AP}$ , em que AP = área da parcela

em m<sup>2</sup>; a tonelada de sacarose por hectare sem palha (TPHsp):

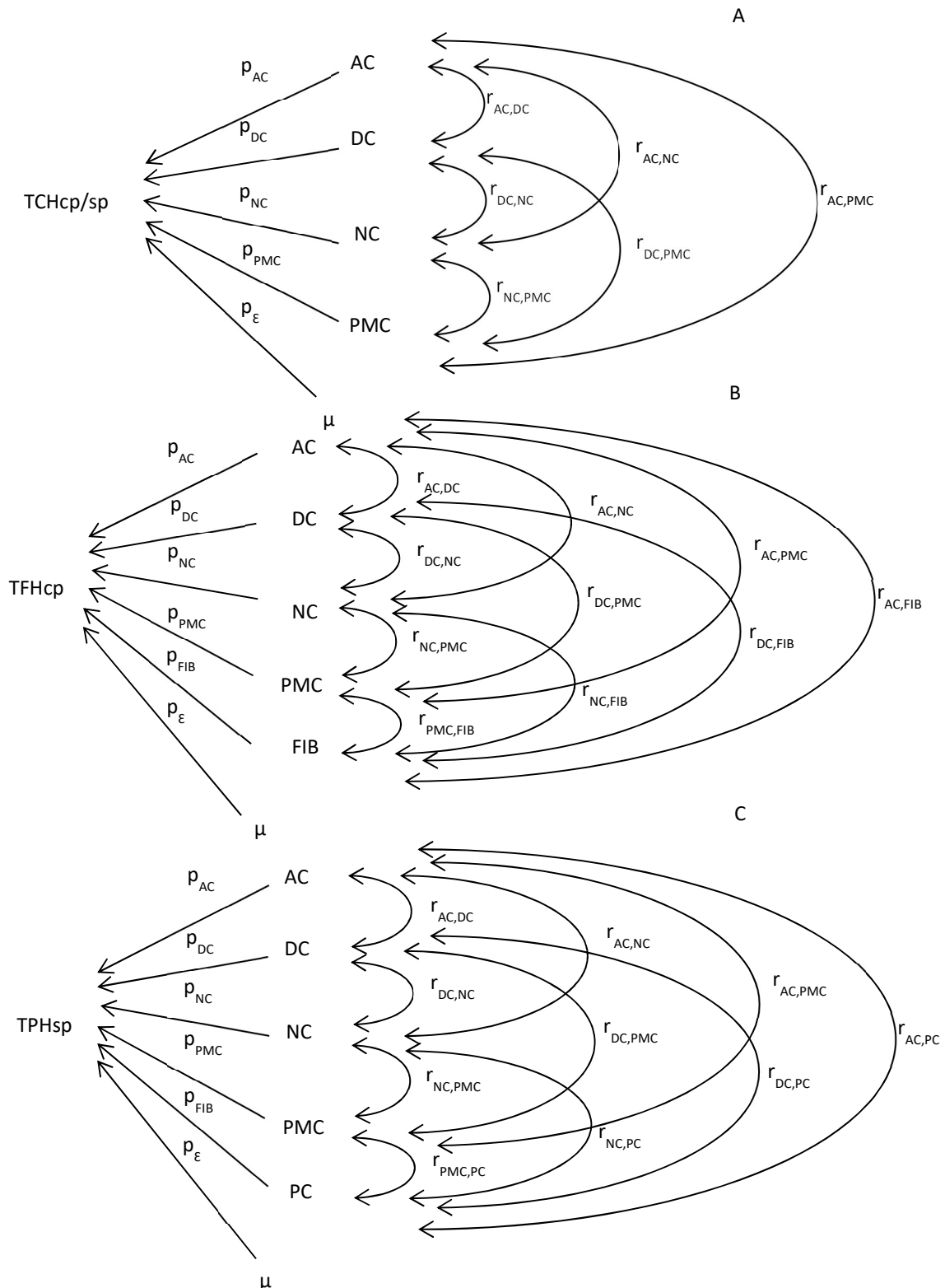
$TPH_{sp} = \frac{(TCH_{sp} \times PC)}{100}$  e a tonelada de fibra por hectare com palha (TFHcp):

$TFH_{cp} = \frac{(TCH_{cp} \times FIB)}{100}$ .

Os dados foram submetidos a análise de modelos mistos via máxima verossimilhança restrita (REML/BLUP) conforme o seguinte modelo estatístico:  $y = Xr + Zf + Wb + Sc + e$ , em que y é o vetor de dados, r é o vetor dos efeitos fixos (repetição) somados à média geral, f é o vetor dos efeitos de famílias de irmãos completos (aleatórios), b é o vetor dos efeitos de blocos incompletos (aleatórios), c é o vetor dos efeitos de clone dentro de família de irmãos completos (aleatórios), e é o vetor de erros ou resíduos (aleatórios). As letras maiúsculas representam as matrizes de incidência para os referidos efeitos (RESENDE, 2007).

A matriz de correlação entre os caracteres foi obtida a partir das médias genéticas, por meio da correlação de Pearson. Em seguida, foi realizado o diagnóstico de multicolinearidade na matriz de correlação entre os caracteres explicativos ( $X'X$ ), de acordo com as recomendações de Montgomery et al. (2012). Posteriormente foi realizado a análise de trilha, conforme os diagramas causais apresentados na Figura 3.1.

FIGURA 3.1 - DIAGRAMAS CAUSAIS ILUSTRANDO O EFEITO DIRETO E INDIRETO DOS CARACTERES DE PRODUÇÃO SOBRE O CARÁCTER PRINCIPAL. A: TONELADA DE COLMOS POR HECTARE COM PALHA E SEM PALHA (TCHcp/sp), B: TONELADA DE FIBRA POR HECTARE COM PALHA (TFHcp), E C: TONELADA DE SACAROSE POR HECTARE SEM PALHA (TPHsp). AS SETAS EM AMBAS AS DIREÇÕES INDICAM ASSOCIAÇÃO MÚTUA ENTRE OS CARACTERES, DETERMINADA PELO COEFICIENTE DE CORRELAÇÃO ( $r_{xy}$ ) E AS SETAS EM ATUAÇÃO DIRETA PARA UM LADO REPRESENTAM O EFEITO DIRETO, DETERMINADO PELO COEFICIENTE DE TRILHA ( $p_{x,y}$ ); AC = ALTURA DO COLMO, DC = DIÂMETRO DO COLMO, NC= NÚMERO DE COLMOS, PMC = PESO MÉDIO DO COLMO, FIB = TEOR DE FIBRA, PC = TEOR DE SACAROSE APARENTE.



As médias genéticas foram obtidas usando o programa Selegen-REML/BLUP (RESENDE, 2016). A matriz de correlação, o diagnóstico de multicolinearidade e as análises de trilha foram realizadas com o auxílio do programa estatístico R (R DEVELOPMENT CORE TEAM, 2016).

### 3.3 RESULTADOS E DISCUSSÃO

As estatísticas descritivas evidenciam a grande amplitude dos dados, como a tonelada de colmos por hectare com palha (TCHcp) em cana planta variando de 13,01 a 315,08 toneladas e a tonelada de fibra por hectare com palha (TFHcp) variando de 7,32 a 44,5 toneladas. E em primeira soqueira a variação da TCHcp foi de 17,33 a 157,57 e para TPHsp variou de 2,26 a 23,71 (Tabela 3.1). A grande variação deve-se a mistura existente na população estudada, com a presença de vários biótipos de cana energia e de diferentes perfis. A ocorrência de falhas na linha propiciou médias reduzidas para TCHcp, TCHsp, TPHsp e TFHcp (Tabela 3.1).

TABELA 3.1 - ESTATÍSTICA DESCRITIVA DOS CARACTERES COMPONENTES DE PRODUÇÃO E DE QUALIDADE EM CANA ENERGIA NOS CICLOS DE CANA-PLANTA E PRIMEIRA SOQUEIRA.

Estatísticas	NC	AC	DC	PMCcp	PMCsp	PC	FIB	TCHcp	TCHsp	TFHcp	TPHsp
Cana-planta											
Média	164,17	3,05	2,09	1,36	1,03	9,01	14,15	141,06	118,55	22,06	10,89
Mínimo	12,90	2,00	1,18	0,54	0,40	2,42	10,67	13,01	43,67	7,32	1,89
Máximo	359,00	4,78	5,71	2,72	2,30	14,17	20,79	315,08	285,57	44,50	32,64
Variância	4332,9	0,13	0,23	0,19	0,14	4,714	3,071	3003,26	1393,6	37,46	22,47
Desvio Padrão	65,825	0,36	0,48	0,43	0,38	2,171	1,752	54,802	37,331	6,121	4,741
Primeira Soqueira											
Média	175,84	2,42	1,97	1,02	0,76	9,66	15,21	112,50	83,53	18,39	8,72
Mínimo	55,00	1,50	1,22	0,40	0,28	5,43	10,29	17,33	12,13	7,92	2,27
Máximo	470,00	3,45	4,08	2,24	1,84	15,78	19,82	257,57	193,78	35,75	23,71
Variância	3682,71	0,13	0,13	0,12	0,08	6,03	4,41	1310,66	816,81	30,59	13,94
Desvio Padrão	60,69	0,36	0,36	0,34	0,28	2,46	2,10	36,20	28,58	5,53	3,73

NC: número de colmos; AC: altura do colmo (m); DC: diâmetro de colmo (cm); PMCcp: peso médio do colmo com palha; PMCsp: peso médio do colmo sem palha; PC: teor de sacarose aparente; FIB: teor de fibra; TCHcp: tonelada de colmos por hectare com palha; TCHsp: tonelada de colmos por hectare sem palha; TFHcp: tonelada de fibra por hectare com palha; TPHsp: tonelada de sacarose por hectare sem palha.

As estimativas dos coeficientes de correlação de TCHcp com TCHsp, TFHcp e TPHsp foram elevadas em cana-planta e primeira soqueira. Para TCHsp as maiores correlações também foram com TCHcp, TPHsp, TFHcp em ambos os ciclos da cultura. Os caracteres de maior correlação com TFHcp em cana-planta foram TCHcp (0,83) e TCHsp (0,79). No ciclo de primeira soqueira as maiores correlações com TFHcp foram TCHcp (0,89) e TCHsp (0,86) (Tabela 3.2).

Os caracteres de maior correlação com TPHsp no ciclo de cana-planta foram TCHsp (0,84), TCHcp (0,69) e PC (0,65), conforme pode ser verificado na tabela 3.2. No ciclo de primeira soqueira os caracteres de maior correlação com TPHsp foram TCHsp (0,75), TCHcp (0,72) e TFHcp (0,61). As correlações elevadas e positivas entre os caracteres avaliados nesse estudo indicam que ambos se correlacionam, e o aumento de um causaria aumento do outro.

TABELA 3.2 - ESTIMATIVAS DOS COEFICIENTES DE CORRELAÇÃO GENÉTICA ENTRE OS CARACTERES. OS VALORES DE CORRELAÇÃO ACIMA DA DIAGONAL PRINCIPAL SÃO REFERENTES AO CICLO DE CANA-PLANTA E ABAIXO DA DIAGONAL PRINCIPAL AO CICLO DE PRIMEIRA SOQUEIRA.

Caracteres	NC	AC	DC	PMCcp	PMCsp	PC	FIB	TCHcp	TCHsp	TFHcp	TPHsp
NC	-	-0,02 <sup>ns</sup>	-0,34 <sup>**</sup>	-0,38 <sup>**</sup>	-0,37 <sup>**</sup>	0,01 <sup>ns</sup>	0,10 <sup>ns</sup>	0,40 <sup>**</sup>	0,31 <sup>**</sup>	0,41 <sup>**</sup>	0,23 <sup>**</sup>
AC	0,06 <sup>ns</sup>	-	0,16 <sup>**</sup>	0,43 <sup>**</sup>	0,42 <sup>**</sup>	0,09 <sup>ns</sup>	0,01 <sup>ns</sup>	0,42 <sup>**</sup>	0,46 <sup>**</sup>	0,46 <sup>**</sup>	0,38 <sup>**</sup>
DC	-0,37 <sup>**</sup>	0,05 <sup>ns</sup>	-	0,57 <sup>**</sup>	0,55 <sup>**</sup>	0,06 <sup>ns</sup>	-0,28 <sup>**</sup>	0,24 <sup>**</sup>	0,30 <sup>**</sup>	0,16 <sup>*</sup>	0,28 <sup>*</sup>
PMCcp	-0,37 <sup>**</sup>	0,07 <sup>ns</sup>	0,74 <sup>**</sup>	-	0,95 <sup>**</sup>	0,15 <sup>*</sup>	-0,27 <sup>**</sup>	0,58 <sup>**</sup>	0,62 <sup>**</sup>	0,48 <sup>**</sup>	0,53 <sup>**</sup>
PMCsp	-0,35 <sup>**</sup>	0,10 <sup>ns</sup>	0,70 <sup>**</sup>	0,95 <sup>**</sup>	-	0,17 <sup>*</sup>	-0,32 <sup>**</sup>	0,54 <sup>**</sup>	0,68 <sup>**</sup>	0,42 <sup>**</sup>	0,60 <sup>**</sup>
PC	-0,09 <sup>ns</sup>	-0,26 <sup>*</sup>	0,06 <sup>ns</sup>	0,06 <sup>ns</sup>	0,02 <sup>ns</sup>	-	-0,25 <sup>**</sup>	0,12 <sup>*</sup>	0,18 <sup>*</sup>	0,03 <sup>ns</sup>	0,65 <sup>**</sup>
FIB	0,14 <sup>ns</sup>	0,23 <sup>ns</sup>	-0,29 <sup>*</sup>	-0,26 <sup>**</sup>	-0,23 <sup>*</sup>	-0,07 <sup>ns</sup>	-	-0,18 <sup>*</sup>	-0,28 <sup>**</sup>	0,23 <sup>**</sup>	-0,34 <sup>**</sup>
TCHcp	0,41 <sup>**</sup>	0,16 <sup>ns</sup>	0,29 <sup>*</sup>	0,57 <sup>**</sup>	0,53 <sup>**</sup>	-0,01 <sup>ns</sup>	-0,12 <sup>ns</sup>	-	0,87 <sup>**</sup>	0,83 <sup>**</sup>	0,69 <sup>**</sup>
TCHsp	0,38 <sup>**</sup>	0,18 <sup>ns</sup>	0,29 <sup>*</sup>	0,56 <sup>**</sup>	0,60 <sup>**</sup>	-0,03 <sup>ns</sup>	-0,11 <sup>ns</sup>	0,97 <sup>**</sup>	-	0,79 <sup>**</sup>	0,84 <sup>**</sup>
TFHcp	0,47 <sup>**</sup>	0,37 <sup>**</sup>	0,15 <sup>ns</sup>	0,42 <sup>**</sup>	0,40 <sup>**</sup>	-0,06 <sup>ns</sup>	0,22 <sup>ns</sup>	0,89 <sup>**</sup>	0,86 <sup>**</sup>	-	0,60 <sup>**</sup>
TPHsp	0,26 <sup>**</sup>	-0,02 <sup>ns</sup>	0,24 <sup>ns</sup>	0,45 <sup>**</sup>	0,49 <sup>**</sup>	0,55 <sup>**</sup>	-0,12 <sup>ns</sup>	0,72 <sup>**</sup>	0,75 <sup>**</sup>	0,61 <sup>**</sup>	-

\*\* correlação significativa a 1% e \* 5% de probabilidade de acordo com o teste t; <sup>ns</sup> correlação não significativa.

NC: número de colmos; AC: altura do colmo; DC: diâmetro do colmo; PMCcp: peso médio do colmo com palha; PMCsp: peso médio do colmo sem palha; PC: teor de sacarose aparente; FIB: teor de fibra; TCHcp: tonelada de colmos por hectare com palha; TCHsp: tonelada de colmos por hectare sem palha; TFHcp: tonelada de fibra por hectare com palha; TPHsp: tonelada de sacarose por hectare sem palha.

Os caracteres TCHcp e TCHsp apresentaram correlações altas e significativas com TFHcp e TPHsp. Logo, a coleta do colmo pode ser tanto com palha quanto sem palha. Sabendo-se que as avaliações fenotípicas em experimentos com cana-de-açúcar são onerosas, principalmente nas fases iniciais pois envolvem um número grande de clones, não há necessidade de realizar a coleta de ambos, mas sim escolher o caracter de fácil obtenção. A palha geralmente é retirada para a fabricação de etanol e açúcar, porém, incrementa a biomassa para a produção de energia.

O NC tanto em cana-planta quanto em primeira soqueira, teve correlação negativa com DC, PMCcp e PMCsp. A correlação negativa é devido à população em estudo, onde um grande número de colmos não está atrelado ao maior peso médio do colmo e a maior competição acaba resultando em diâmetro menor. Em cana energia os materiais apresentam muitos colmos, porém finos, o que resulta em menor peso médio do colmo.

A obtenção de genótipos com alto teor de fibra, característica almejada em clones de cana energia, envolve hibridações entre acessos de *Saccharum spontaneum* e *Saccharum robustum* com as variedades modernas de cana-de-açúcar (*Saccharum* spp.), em razão da mesma apresentar alto teor de fibra. Os descendentes desses cruzamentos envolvendo os dois genitores, possuem grande capacidade de perfilhamento e produção de biomassa (KIM e DAY, 2011; SILVEIRA et al., 2015b).

O DC apresentou correlação negativa de (-0,28) com fibra no ciclo de cana-planta e (-0,29) em primeira soqueira (Tabela 3.2), dessa forma, um menor diâmetro de colmo resultaria em maior teor de fibra. Como a fibra se concentra na casca, quanto menor o diâmetro, maior a razão da área de superfície para volume. Milligan et al. (1990) constataram que o número de colmos e o diâmetro são inversamente relacionados, resultado que também foi encontrado no presente trabalho.

Os estudos de correlação tem importância para quantificação da magnitude e direção existente entre dois caracteres (TENA, 2016; FALCONER e MACKAY, 1996). Porém, deve-se tomar cuidado, pois a correlação entre dois caracteres pode não estar associada a uma medida de causa e efeito, e a interpretação direta pode resultar em erros no processo de seleção. Quando dois caracteres apresentam correlação alta, isso pode ser resultado da influência de um destes ou de um conjunto de caracteres (SILVA et al., 2009).

A análise de trilha permite obter a relação de causa e efeito entre dois caracteres. Assim, as estimativas dos efeitos diretos e indiretos dos caracteres componentes da produção sobre TCHcp, TCHsp, TFHcp e TPHsp foram obtidas via análise de trilha. As matrizes de correlação entre os caracteres explicativos não apresentaram problemas de multicolinearidade, número de condição ( $N_c < 100$ ) (Tabelas 3.3 e 3.4), segundo critérios de Montgomery et al. (2012).

Os coeficientes de correlação entre os caracteres explicativos em todas as análises de trilha realizadas, foram decompostos em efeitos diretos e indiretos. Os valores elevados dos coeficientes de determinação ( $R^2$ ) e o pequeno efeito residual, indicam uma eficiente contribuição dos caracteres explicativos em relação aos caracteres dependentes (Tabelas 3.3 e 3.4).

TABELA 3.3 - EFEITOS DIRETOS E INDIRETOS DOS CARACTERES EXPLICATIVOS SOBRE OS CARACTERES PRINCIPAIS TONELADA DE COLMOS POR HECTARE COM PALHA (TCHcp) E TONELADA DE COLMO POR HECTARE SEM PALHA (TCHsp) EM CANA-PLANTA E PRIMEIRA SOQUEIRA NA SEGUNDA FASE DE TESTE.

Efeitos	Caracteres Explicativos Cana-Planta			
	NC	AC	DC	PMCcp
Direto TCHcp	0,713	0,094	0,025	0,791
Indireto via NC	-	-0,012	-0,243	-0,270
Indireto via AC	-0,002	-	0,015	0,039
Indireto via DC	-0,008	0,004	-	0,014
Indireto via PMCcp	-0,299	0,336	0,447	-
R <sup>2</sup>		0,788		
$\hat{p}_\varepsilon$		0,460		
Nc		5,783		
Efeitos	Caracteres Explicativos Primeira soqueira			
	NC	AC	DC	PMCcp
Direto TCHcp	0,700	0,058	-0,133	0,918
Indireto via NC	-	0,044	-0,262	-0,258
Indireto via AC	0,004	-	0,003	0,004
Indireto via DC	0,049	-0,006	-	-0,098
Indireto via PMCcp	-0,338	0,065	0,681	-
R <sup>2</sup>		0,782		
$\hat{p}_\varepsilon$		0,467		
Nc		7,813		
Efeitos	Caracteres Explicativos Cana-Planta			
	NC	AC	DC	PMCSp
Direto TCHsp	0,639	0,100	0,033	0,859
Indireto via NC	-	-0,011	-0,218	-0,237
Indireto via AC	-0,001	-	0,016	0,042
Indireto via DC	-0,011	0,005	-	0,018
Indireto via PMCSp	-0,319	0,363	0,473	-
R <sup>2</sup>		0,839		
$\hat{p}_\varepsilon$		0,400		
Nc		5,502		
Efeitos	Caracteres Explicativos Primeira soqueira			
	NC	AC	DC	PMCSp
Direto TCHsp	0,661	0,056	-0,066	0,868
Indireto via NC	-	0,041	-0,247	-0,232
Indireto via AC	0,004	-	0,003	0,006
Indireto via DC	0,024	-0,003	-	-0,046
Indireto via PMCSp	-0,305	0,087	0,604	-
R <sup>2</sup>		0,763		
$\hat{p}_\varepsilon$		0,486		
Nc		6,508		

NC: número de colmo, AC: altura de colmo, DC: diâmetro de colmo, PMCcp: peso médio do colmo com palha, PMCSp: peso médio do colmo sem palha, coeficiente de determinação (R<sup>2</sup>), efeito da variável residual ( $\hat{p}_\varepsilon$ ), número de condição (Nc).

TABELA 3.4 - EFEITOS DIRETOS E INDIRETOS DOS CARACTERES EXPLICATIVOS SOBRE OS CARACTERES PRINCIPAIS TONELADA DE FIBRA POR HECTARE COM PALHA (TFHcp) E TONELADA DE SACAROSE POR HECTARE SEM PALHA (TPHsp) EM CANA-PLANTA E PRIMEIRA SOQUEIRA NA SEGUNDA FASE DE TESTE.

Efeitos	Caracteres Explicativos Cana-Planta				
	NC	AC	DC	PMCcp	FIB
Direto TFHcp	0,676	0,144	0,051	0,746	0,377
Indireto via NC	-	-0,011	-0,231	-0,256	0,068
Indireto via AC	-0,002	-	0,023	0,061	0,001
Indireto via DC	-0,017	0,008	-	0,028	-0,014
Indireto via PMCcp	-0,283	0,317	0,422	-	-0,202
Indireto via FIB	0,038	0,004	-0,107	-0,102	-
R <sup>2</sup>			0,797		
$\hat{p}_\varepsilon$			0,450		
Nc			6,255		
Efeitos	Caracteres Explicativos Primeira soqueira				
	NC	AC	DC	PMCcp	FIB
Direto TFHcp	0,671	0,225	-0,165	0,839	0,238
Indireto via NC	-	0,042	-0,251	-0,247	0,095
Indireto via AC	0,014	-	0,012	0,016	0,051
Indireto via DC	0,062	-0,008	-	-0,122	0,048
Indireto via PMCcp	-0,309	0,060	0,622	-	-0,217
Indireto via FIB	0,034	0,054	-0,069	-0,061	-
R <sup>2</sup>			0,782		
$\hat{p}_\varepsilon$			0,466		
Nc			8,408		
Efeitos	Caracteres Explicativos Cana-Planta				
	NC	AC	DC	PMCsp	PC
Direto TPHsp	0,470	0,064	0,048	0,629	0,526
Indireto via NC	-	-0,008	-0,161	-0,175	0,006
Indireto via AC	-0,001	-	0,010	0,027	0,005
Indireto via DC	-0,016	0,007	-	0,026	0,003
Indireto via PMCsp	-0,234	0,265	0,347	-	0,107
Indireto via PC	0,007	0,048	0,031	0,090	-
R <sup>2</sup>			0,861		
$\hat{p}_\varepsilon$			0,372		
Nc			5,703		
Efeitos	Caracteres Explicativos Primeira soqueira				
	NC	AC	DC	PMCsp	PC
Direto TPHsp	0,522	0,039	-0,128	0,748	0,600
Indireto via NC	-	0,032	-0,195	-0,183	-0,045
Indireto via AC	0,003	-	0,002	0,004	-0,010
Indireto via DC	0,047	-0,006	-	-0,089	-0,008
Indireto via PMCsp	-0,262	0,074	0,519	-	0,017
Indireto via PC	-0,052	-0,156	0,037	0,014	-
R <sup>2</sup>			0,805		
$\hat{p}_\varepsilon$			0,442		
Nc			6,528		

NC: número de colmos, AC: altura de colmo, DC: diâmetro de colmo, PMCcp: peso médio do colmo com palha, PMCsp: peso médio do colmo sem palha, FIB: teor de fibra, PC: teor de sacarose aparente, coeficiente de determinação (R<sup>2</sup>), efeito da variável residual ( $\hat{p}_\varepsilon$ ), número de condição (Nc).

Os caracteres que mais contribuíram para explicar TCHcp foram o PMCcp e NC que apresentaram valores de efeito direto de (0,791), (0,713) para o ciclo de cana-planta e (0,918) e (0,700) para o ciclo de primeira soqueira, respectivamente.

Enquanto que, os caracteres de maior efeito direto com TCHsp em ambos os ciclos foram PMCsp seguido do NC. Os demais caracteres, AC e DC tiveram efeitos diretos inferiores aos efeitos residuais, demonstrando pouca importância no processo de seleção para TCHcp e TCHsp em cana energia (Tabela 3.3).

Os trabalhos de James (1971) e Reddy e Reddi (1986) demonstraram contribuição semelhante do número de colmos e diâmetro do colmo sobre tonelada de colmos por hectare, seguidos de altura do colmo. Sukhchain e Saini (1997) também obtiveram alto efeito direto de NC sobre TCH em cana convencional. Espósito et al. (2012), Brasileiro et al. (2013) e Silva et al. (2009) também obtiveram altos efeitos diretos de NC sobre a produtividade de colmos em cana convencional.

Silveira et al. (2015b) mostraram que a seleção das famílias de cana energia na primeira fase de teste (T1) com as maiores médias para TCH, pode ser realizada via número de colmos e peso médio do colmo, pois esses dois caracteres foram os principais responsáveis pela determinação da produtividade de colmos (TCH). No presente trabalho, conduzido na segunda fase de teste (T2), onde a seleção passa a ser clonal, os caracteres PMCcp, PMCsp e NC se mantiveram como os de maior efeito sobre a produtividade, demonstrando que ambos têm grande efeito sobre a tonelada de colmos por hectare.

O PMCcp e o NC foram os caracteres de maior efeito direto sobre TFHcp, apresentando efeitos diretos de 0,746, 0,676 para o ciclo de cana-planta e de 0,839, 0,671 para o ciclo de primeira soqueira, respectivamente. O PMCsp, PC e NC foram os caracteres de maior efeito direto sobre a TPHsp em ambos os ciclos (Tabela 3.4). Os caracteres AC e DC tiveram efeitos diretos inferiores aos efeitos residuais, demonstrando ter pouca importância no processo de seleção para TFHcp e TPHsp de cana energia. Silveira et al. (2015b) ao avaliarem famílias de cana energia, observaram que o número de colmos foi o segundo caractere de maior efeito direto na tonelada de fibra por hectare e na tonelada de sacarose por hectare em cana energia.

Em ambos os ciclos, os caracteres que mais contribuíram para explicar TCHcp, TCHsp TFHcp, TPHsp foram PMCcp, PMCsp seguido de NC (Tabelas 3.3 e 3.4), indicando que tanto para cana-planta, quanto para primeira soqueira a seleção pode ser baseada em um dos dois, ou seja, no NC ou no PMC, pois ambos se correlacionam de forma negativa. Logo o uso dos dois caracteres para a seleção não resultaria em ganhos para TCHcp, TCHsp, TFHcp e TPHcp (Tabela 3.2).

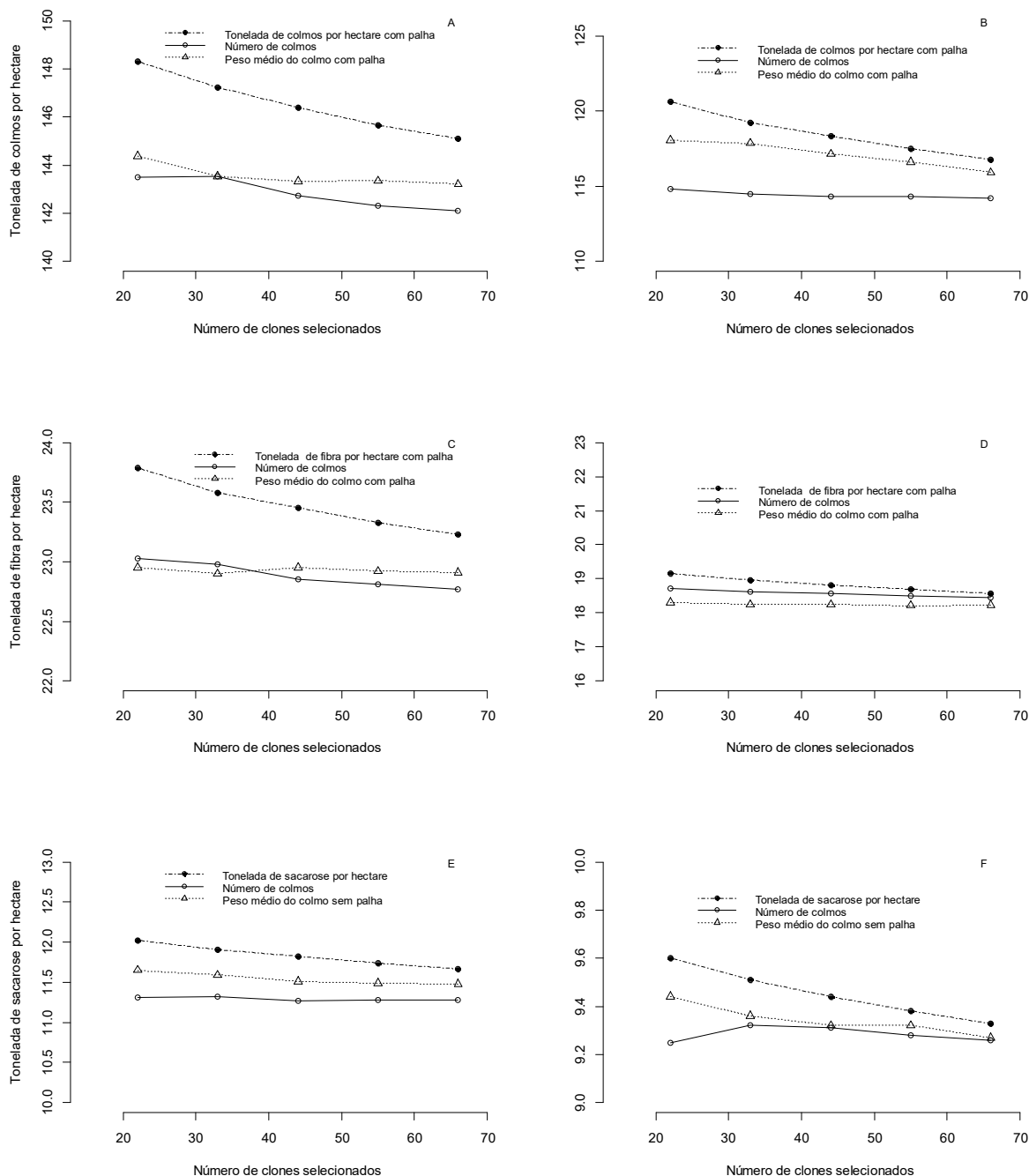
Santchurn et al. (2012) mostraram que existe correlação negativa entre fibra e açúcar, resultado também encontrado no presente trabalho. Os mesmos autores ainda enfatizam a dificuldade de selecionar materiais para ambos os caracteres. No entanto, Ramos et al. (2017) mostraram que é possível manter o teor de sacarose das atuais variedades comerciais e elevar o teor de fibra das futuras variedades.

A cana-de-açúcar por ser uma cultura perene, pode passar por dois momentos distintos, sendo um período chuvoso, onde ocorre o acúmulo de fibra e um período de estresse hídrico, onde cessa o crescimento vegetativo e ocorre o acúmulo de sacarose (KENNEDY, 2008). É possível que o efeito dos genes em dividir fotoassimilados para maior crescimento do colmo ou para maior armazenamento, tenderia a contribuir para uma correlação genética negativa entre o crescimento e o armazenamento (RAMOS et al., 2017).

Foi realizado uma simulação, utilizando os melhores clones para tonelada de colmos por hectare (TCH), tonelada de fibra (TFH) e tonelada de sacarose (TPH) levando em consideração uma seleção direta via TCH, TFH e TPH, e uma seleção indireta via número de colmos (NC) e peso médio do colmo com palha e sem palha (PMCcp e PMCsp). A média para tonelada de colmos por hectare é superior com a seleção direta quando comparado com uma seleção indireta via NC ou PMCcp . Porém, a medida que aumenta o número de clones selecionados as médias de TCH via seleção direta e indireta se aproximam, para os ciclos de cana-planta e primeira soqueira (Figura 3.2 A e B).

A média em relação aos melhores clones selecionados para tonelada de fibra por hectare e tonelada de sacarose por hectare se aproximam quando aumenta o número de clones selecionados, tanto com utilização de seleção direta quanto seleção indireta, isso para o ciclo de cana-planta e primeira soqueira (Figura 3.2 C, D, E, F). A média para tonelada de fibra e tonelada de sacarose no ciclo de primeira soqueira ao utilizar seleção direta via número de colmos é maior, se equivalendo para tonelada de sacarose por hectare a medida que aumenta o número de clones selecionados (Figura 3.2 D e F).

FIGURA 3.2 - MÉDIAS PARA TONELADA DE COLMOS POR HECTARE, TONELADA DE FIBRA POR HECTARE E TONELADA DE SACAROSE POR HECTARE NOS CICLOS DE CANA-PLANTA (A, C, E) E PRIMEIRA SOQUEIRA (B, D, F) EM FUNÇÃO DA SELEÇÃO DIRETA OU INDIRETA E DO NÚMERO DE CLONES SELECIONADOS.



A utilização de seleção indireta para clones de cana energia utilizando o PMCcp, PMCsp ou NC, que foram os caracteres de maior efeito direto com TCHcp, TFHcp e TPHsp, acaba não sendo tão vantajosa para TCH quanto maior a taxa de

seleção. A necessidade de pesar os colmos, para a aplicação da seleção indireta via PMCcp, faz com que a escolha de seleção direta com base na própria TCH seja mais vantajosa.

A seleção para TFHcp e TPHsp pode ser realizada via NC, nas situações onde precisa-se avaliar muitos clones. O NC permitiu a identificação de clones com médias similares as obtidas com a seleção direta via TFHcp e TPHs, principalmente quando a taxa de seleção é elevada, além de ser um caractere de fácil mensuração, a seleção via NC torna o processo menos oneroso.

Os resultados observados neste trabalho indicam que, se os programas de melhoramento genético forem voltados para obtenção de clones para maior teor de fibra, a seleção deve ser direcionada para clones com maior número de colmos, o que resultará em maior produtividade de biomassa.

### 3.4 CONCLUSÕES

Os caracteres peso médio do colmo com palha ou peso médio do colmo sem palha e o número de colmos, foram os componentes de maior efeito direto sobre a produtividade de cana, fibra e açúcar nos ciclos de cana-planta e primeira soqueira.

### 3.5 REFERÊNCIAS

APARECIDO, L.E.O.; ROLIM, G.S.; RICHETTI, J.; SOUZA, P.S.; JOHANN, J.A. Köppen, Thornthwaite and Camargo climate classifications for climatic zoning in the State of Paraná, Brazil. **Ciência e Agrotecnologia**, v. 40, n.4, p. 405-417, 2016.

BAFFA, D.C.; COSTA, P.M.A.; SILVEIRA, G.; LOPES, F.J.F.; BARBOSA, M.H.P.; LOUREIRO, M.E.; CRUZ, C.D.; PETERNELLI, L.A. Path analysis for selection of Saccharification-Efficient sugarcane genotypes through agronomic traits. **Agronomy Journal**, v. 106, n.5, 2014.

BARBOSA, M.H.P.; SILVEIRA, L.C.I. Breeding and cultivar recommendations. In Santos F, Borém A and Caldas C (eds.) **Sugarcane: bioenergy, sugar and ethanol: technology and prospects**. MAPA/ACS: UFV/DEA, Brasília, p. 301-318, 2012.

BELLO, O.B., S.Y. ABDULMALIQ, M.S. AFOLABI.; S.A. IGE. Correlation and path coefficient analysis of yield and agronomic characters among open pollinated maize varieties and their F1 hybrids in a diallel cross. **African Journal Biotechnology**. v.9, p.2633-2639, 2010.

BRASILEIRO, B.P.; PETERNELLI, L.A.; BARBOSA, M.H.P. Consistency of the results of path analysis among sugarcane experiments. **Crop Breeding and Applied Biotechnology**, v.13, p.99-107, 2013.

CHAUDHARY, R.R.; JOSHI, B.K. Correlation and Path Coefficient Analyses in Sugarcane. **Nepal Agriculture Research Journal**, v. 6, 2005.

ESPÓSITO, D.P.; PETERNELLI, L.A.; DE PAULA, T.O.M.; BARBOSA, M.H.P. Path analysis using phenotypic and genotypic values for yield components in the selection of sugarcane families. **Ciência Rural**, v.42, n.1, p.38-44, 2012.

FALCONER, D.S.; MACKAY, T.F.C. **Introduction to qualitative genetics**. 4.ed. New York, Longman. p.464, 1996.

FERNANDES, A.C. **Cálculos na Agroindústria da cana-de-açúcar**. STAB, 3ª.ed., 416p, 2011.

FERREIRA, F.M.; BARROS, W.S.; SILVA, F.L.; BARBOSA, M.H.P.; CRUZ, C.D.; BASTOS, I.T. Relação fenotípica e genotípica entre componentes de produção em cana-de-açúcar. **Bragantia**, v.66, n.4, p.605-610, 2007.

JAMES, N.I. Yield components in random and selected sugarcane populations. **Crop Science**, v.11, p.906-908, 1971.

KENNEDY, A.J. Prospects for combining high sucrose content with increased fibre to generate multi-purpose cane varieties. In: **Proceeding of the Conference of West Indies Sugar Technologists**, Jamaica, 2008.

KIM, M.; DAY, D.F. Composition of sugar cane, energy cane, and sweet sorghum suitable for ethanol production at Louisiana sugar mills. **Journal of Industrial Microbiology and Biotechnology**, v. 38, n. 7, p. 803-807, 2011.

LOUREIRO, M.E.; BARBOSA, M.H.P.; LOPES, F.J.F.; SILVÉRIO, F.O. Sugarcane breeding and selection for more efficient biomass conversion in cellulosic ethanol. In: Buckeridge, M.S,Goldman, G.H (eds) **Routes to cellulosic ethanol**. (p.199-239). New York Dordrecht Heidelberg London: Springer, 2011.

MARCHEZAN, E.; MARTIN, T.N.; SANTOS, F.M. DOS.; CAMARGO, E.R. Análise de coeficiente de trilha para os componentes de produção em arroz. **Ciência Rural**, v.35, n.5, p.1027-1033, 2005.

MILLIGAN, S.B.; GRAVOIS, K.A.; BISCHOFF, K.P.; MARTIN, F.A. Crops effects on broad-sense heritabilities and genetic variances of sugarcane yield components. **Crop Science**. v.30, p.344-349, 1990.

MONTGOMERY, D.C.; PECK, E.A.; VINING, G.G. **Introduction to linear regression analyses**. Wiley, p. 504, 2012.

NARDINO, M.; SOUZA, V.Q.; BARETTA, V.A. KONFLANZ.; CARVALHO, I.R.; FOLLMANN, D.N.; CARON, B.O. Association of secondary traits with yield in maize F1's. **Ciência Rural** v.46, p.776-782, 2016.

OLIVOTO, T.; SOUZA, V.Q de.; NARDINO, M.; CARVALHO, I.R.; FERRARI, M.; PELEGRIN, A.J de.; SZARESKI, V.J.; SCHMIDT, D. Multicollinearity in Path Analysis: A Simple Method to Reduce Its Effects. **Agronomy Journal**, v.109, n.1, 2017.

R DEVELOPMENT CORE TEAM. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org>, 2016.

RAMOS, R.S.; BRASILEIRO, B.P; SILVEIRA, L.C.I.; KIST, V.; PETERNELLI, L.A.; BARBOSA, M. H. P. Selection of parents, families and clones to obtain energy cane. **Agronomy Journal**, v.109, p.1-7, 2017.

REDDY, C.R.; REDDI, M.V. Degree of genetic determination, correlation and genotypic and phenotypic path analysis of cane and sugar yield in sugarcane. **Indian Journal of Genetics and Plant Breeding**, v.46, p.550-557, 1986.

RESENDE, M.D.V. **SELEGEN-REML/BLUP Sistema Estatístico e seleção genética Computadorizada Via Modelos Lineares Mistos**. Ed Embrapa, 1ª ed, 360p, 2007.

RESENDE, M.D.V. Software Selegen-REML/BLUP: a useful tool for plant breeding. **Crop Breeding and Applied Biotechnology** v.16, p.330-339, 2016.

SANTCHURN, D.; RAMDOYAL, K.; BADALOO, M.G.H.; LABUSCHAGNE, M. From sugar industry to cane industry: investigations on multivariate data analysis techniques in the identification of different high biomass sugarcane varieties. **Euphytica**, v.185, p.543-558, 2012.

SANTOS, H.G.; JACOMINE, P.K.T.; ANJOS, L.H.C.; OLIVEIRA, V.A.; LUBRERAS, J.F.; COELHO, M.R.; ALMEIDA, J.A.; CUNHA, T.J.F.; OLIVEIRA, J.B. **Sistema Brasileiro de Classificação de solos**. 3.ed. Brasília, EMBRAPA, 353p, 2013.

SILVA, F.L.; PEDROZO, C.A.; BARBOSA, M.H.P.; RESENDE, M.D.V.; PETERNELLI, L.A.; COSTA, P.M DE A.; VIEIRA, M.S. Análise de trilha para os componentes de produção de cana-de-açúcar via blup. **Revista Ceres**, v.56, n.3, p.308-314, 2009.

SILVEIRA, L.C.I.; BRASILEIRO, B.P.; KIST, V.; WEBER, H.; DAROS, E.; PETERNELLI, L.A.; BARBOSA, M.H.P. Selection strategy in families of energy cane based on biomass production and quality traits. **Euphytica**, v.204, n.2, p.443-455, 2015a.

SILVEIRA, L.C.I.; BRASILEIRO, B.P.; KIST, V.; DAROS, E.; PETERNELLI, L.A. Genetic diversity and coefficient of kinship among potential genitors for obtaining cultivars of energy cane. **Revista Ciência Agronômica**, v. 46, n. 2, p. 358-368, 2015b.

SUKHCHAIN, S.D.; SAINI, G.S. Inter-relationships among cane yield and commercial cane sugar and their component traits in autumn plant crop of sugarcane. **Euphytica**, v.95, p. 109-114, 1997.

TENA, E.; MEKBIB, F.; AYANA, A. Correlation and Path Coefficient Analyses in Sugarcane Genotypes of Ethiopia. **American Journal of Plant Sciences**, v.7, n.10, p.1490-1497, 2016.

## 4 CAPÍTULO II - SELEÇÃO DE CLONES DE CANA ENERGIA VIA MODELO LOGÍSTICO

### RESUMO

O processo de seleção de clones de cana-de-açúcar é considerado subjetivo, principalmente nas primeiras fases de seleção (T1 e T2) onde a seleção ocorre de forma massal pois o número reduzido de material propagativo na segunda fase de seleção não possibilita o uso de repetições, tornando o experimento com reduzida precisão experimental. O trabalho teve por objetivo identificar os caracteres agrônômicos de maior importância durante o processo de seleção clonal na segunda fase de teste em cana energia e avaliar a eficiência do modelo logístico em prever os genótipos a serem selecionados. Foram realizadas avaliações em 220 clones no ciclo de cana-soca, os caracteres avaliados foram: número de colmos, diâmetro do colmo, altura do colmo, peso médio do colmo, teor de sacarose aparente, teor de fibra. Também foram atribuídas notas para os caracteres despalha, diâmetro do colmo, altura do colmo, número de colmos, presença de brotação lateral, saliência da gema, hábito de crescimento, floração, chochamento, nota para vigor, nota para incidência de carvão, notas para severidade de ferrugem marrom e ferrugem alaranjada. Os dados foram submetidos a análise de regressão logística binária. O número de colmos é um caractere importante para a seleção de clones de cana energia, as plantas que obtiveram nota elevada para esse caractere tiveram maior chance de serem selecionadas, bem como plantas que obtiveram nota menor para carvão tiveram maior chance de serem selecionadas. A utilização do método massal selecionou 38 clones no T2, enquanto os modelos ajustados por regressão logística, denominados de qualitativo e quantitativo, selecionaram 42 e 22 clones respectivamente. A capacidade preditiva dos modelos, qualitativo e quantitativo foram 94% e 87% respectivamente. A utilização de um modelo qualitativo se mostrou eficiente para prever os genótipos de cana energia a serem selecionados, podendo ser utilizado no processo de seleção clonal.

Palavras chave: *Saccharum* spp, melhoramento genético, biomassa.

## SELECTION OF ENERGY CANE CLONES BY LOGISTIC MODEL

### ABSTRACT

The selection process of sugarcane clones is considered to be subjective, especially in the first phases where selection occurs massively and the reduced number of propagating material in the second test phase does not allow for repeatability making the experiment with reduced experimental precision. The objective of this work was to identify the most important agronomic traits during the clonal selection process in the second test phase in sugar cane and to evaluate the efficiency of the logistic model in predicting the number of genotypes to be selected. Evaluations were carried out on 220 clones in the first ratoon. The number of stalks, stalk diameter, stem height, stalk weight, number of harvests / plant, apparent sucrose content and fiber content were evaluated. It was assigned notes for straw removal, stalk diameter, stalk number, stalk height, lateral sprouting, growth habit, flowering, stalk without broth, general note, note for coal diseases, brown rust and orange rust. The data were submitted to binary logistic regression analysis. The stalk number is an important feature for the selection of energy cane clones, the plants that obtain a higher grade for this character have a greater chance of being selected, as well as, plants that obtain a lower grade for smut will have a greater chance of being selected. The use of the mass method selected 38 clones in T2, while the logistic regression models, called qualitative and quantitative, selected 42 and 22 clones respectively. The predictive capacity of the models, qualitative and quantitative were 94% and 87%, respectively. The use of a qualitative model proved to be efficient to predict the number of sugar cane genotypes to be selected and could be used as a selection strategy.

Keywords: genetic improvement, biomass, *Saccharum* spp.

## 4.1 INTRODUÇÃO

A seleção individual nos programas de melhoramento genético é realizada em todas as fases (SKINNER et al., 1987). A primeira fase de seleção envolve a avaliação de clones que foram plantados a partir de sementes verdadeiras, a segunda etapa de seleção, ocorre quando os clones, selecionados na primeira fase de teste são propagados, em parcelas sem repetições. A maior eficiência de seleção dos clones nas fases T1 e T2 ainda é um desafio para obtenção de variedades de cana-de-açúcar (ZHOU et al., 2012).

Nas primeiras fases do melhoramento genético da cana-de-açúcar tem-se uma baixa precisão experimental devido ao pouco material propagativo. Nas fases posteriores, à medida que se aumenta a quantidade de material propagativo, aumenta-se o número de repetições e consequentemente, a precisão experimental e a acurácia de seleção.

O método massal, que leva em consideração o fenótipo dos indivíduos, tem sido utilizado nas fases iniciais para selecionar clones produtivos (KIMBENG; COX, 2003). A seleção visual pode ser considerada como subjetiva (HOGARTH e BERDING, 2006; BARBOSA et al., 2005), principalmente nas primeiras fases, em razão dos indivíduos serem propagados sem repetição e do efeito da interação genótipo ambiente (MCRAE e JACKSON, 1998).

Na seleção visual o descarte de um genótipo envolve a combinação de diferentes caracteres como a ocorrência de floração, a presença de doenças, além do conteúdo de açúcar e fibra que também podem influenciar na decisão (BRASILEIRO et al., 2016). O conhecimento de quais caracteres tem maior importância na seleção dos clones de cana energia, possibilita auxiliar nos programas de melhoramento genético, diminuindo o número de caracteres a serem avaliados.

A análise de regressão logística foi desenvolvida para casos em que a variável resposta é categórica (binário ou multinomial). A escolha de selecionar ou rejeitar um clone é uma resposta binária, logo, uma análise de regressão logística pode ser aplicada para avaliar o efeito ou poder de discriminação de cada caractere agrônomo utilizado como critério de seleção, bem como a predição de genótipos a serem selecionados (AGRESTI, 2007).

O uso do modelo logístico como ferramenta estatística de apoio à seleção, com base em componentes de produção, tem se mostrado eficiente em cana-de-açúcar (ZHOU et al., 2014; BRASILEIRO et al., 2015; BRASILEIRO et al., 2016), porém, não se tem trabalhos utilizando essa técnica para a seleção de clones de cana energia.

Portanto, o objetivo desse trabalho é identificar os caracteres agronômicos de maior importância durante o processo de seleção clonal na segunda fase de teste em cana energia e avaliar a eficiência do modelo logístico em prever os genótipos a serem selecionados.

## 4.2 MATERIAL E MÉTODOS

O trabalho foi desenvolvido na Estação Experimental de Paranaíba, pertencente a Universidade Federal do Paraná, com latitude de 23°05'S; longitude de 52°27'W; altitude de 503 m. O clima da região é classificado como Cfa segundo Köppen, com precipitação anual de 1200 a 1400 mm, temperatura média anual entre 22 e 23°C (APARECIDO et al., 2016) e solo Latossolo Vermelho distrófico (SANTOS et al., 2013).

No experimento com 220 clones selecionados na primeira fase de seleção, que foi implantado em junho de 2014, em delineamento de blocos incompletos e cada bloco foi composto por 22 clones. As parcelas foram constituídas de duas linhas de 5 m, espaçadas em 1,4 m e com plantio de 18 gemas por metro. As avaliações foram realizadas no final do ciclo de primeira soqueira, em julho de 2016.

Os caracteres avaliados foram: diâmetro do colmo (DC), com a medição em centímetros do terço médio de 10 colmos por parcela; número de colmos (NC), a partir da contagem dos colmos presentes em um metro dentro da parcela; altura do colmo (AC), com a medição em metros de 10 colmos por parcela a partir do nível do solo até a primeira aurícula visível; peso médio do colmo (PMC), com a pesagem em quilogramas de 10 colmos por parcela; incidência de carvão (CARV), realizada a partir da contagem do número de colmos afetados; teor de sacarose aparente (PC) e o teor de fibra (FIB). Os dois últimos caracteres foram obtidos a partir da análise tecnológica, seguindo as normas técnicas de determinação da qualidade da matéria prima, conforme recomendado por Fernandes (2011).

As avaliações visuais foram feitas utilizando atribuição de notas para os caracteres: despalha (DES), que consiste na soltura da bainha; diâmetro do colmo (DCn); altura do colmo (ACn); número de colmos (NCn); presença de brotação lateral (PB); saliência da gema (SG); hábito de crescimento (HAB); florescimento (FLO); chochamento (CHO): considerando o ponto médio do colmo; nota para vigor (NG): foi atribuída aos 9 meses; ferrugem marrom (FER); ferrugem alaranjada (FAL) e incidência de carvão (CARVn). As notas atribuídas para os caracteres estão representadas na tabela 4.1 e 4.2.

TABELA 4.1 - DESCRIÇÃO DOS CARACTERES AVALIADOS E SUAS RESPECTIVAS NOTAS EM CANA ENERGIA. FONTE: RIDESA.

Caracteres	Notas				
	1	2	3	4	5
NG*	Excelente	Ótimo	Bom	Regular	Péssimo
CHO	Ausente	¼ do Ø	1/2 do Ø	> 50% Ø	
PB	Ausente	Pouco	Regular	Abundante	
FLO	Ausente	velinha	panícula		
DES	Fácil	Médio	Difícil		
DCn	Fino (<2,0cm)	Médio (2,0-3cm)	Grosso (3-3,5cm)		
ACn	Baixo	Médio	Alto		
NCn	Pouco	Médio	Elevado		
SG	2mm	3mm	4mm		
HAB	Ereto	Semi-ereto	Decumbente		

\*NG: nota para vigor; CHO: chochamento; PB: presença de brotação; FLO: florescimento, DES: despalha, DCn: diâmetro do colmo, ACn: altura do colmo, NCn: número de colmos, SG: saliência da gema, HAB: hábito de crescimento.

TABELA 4.2 - NOTAS ATRIBUÍDAS PARA PRESENÇA DE FERRUGEM MARROM, FERRUGEM ALARANJADA E INCIDÊNCIA DE CARVÃO EM CANA ENERGIA.

Notas	caracteres		
	FER	FAL (%)	CARVn
1	Altamente resistente	0,06	0
2	Resistente	0,14	1-15
3	Moderadamente resistente	0,36	15-30
4	Moderadamente suscetível	0,89	30-45
5	Suscetível	2,17	45-60
6	Altamente suscetível	5,18	>60
7		11,87	
8		24,9	
9		45	

FER: ferrugem marrom; FAL: ferrugem alaranjada; CARVn: incidência de carvão

As doenças ferrugem marrom (FER) e ferrugem alaranjada (FAL) foram avaliadas utilizando escalas diagramáticas desenvolvidas por Amorin et al. (1987) e Klosowski et al. (2013) respectivamente. Na avaliação da incidência de carvão foi considerado todas as plantas da parcela, sendo contado o número de colmos com

incidência da doença e posteriormente atribuindo notas de 1 a 5. As doenças foram avaliadas aos 9 meses em primeira soqueira, em março de 2016.

A atribuição das notas para os caracteres avaliados e a seleção massal foram realizadas por um técnico do programa de melhoramento genético da cana-de-açúcar com experiência na seleção visual. A variável resposta consiste em selecionar ou não selecionar um clone, recebendo classificação um os clones que foram selecionados e classificação zero os clones que foram rejeitados no processo de seleção realizado a campo.

Os dados foram submetidos a análise de regressão logística binária com a utilização do programa estatístico R (R DEVELOPMENT CORE TEAM, 2016), na predição do modelo se utilizou ponto de corte de 0,5, dessa forma selecionando os indivíduos que apresentavam uma probabilidade de seleção acima de 0,5. As estimativas dos parâmetros dos coeficientes das variáveis preditoras estimadas pela regressão logística foram utilizadas para construir as funções de distribuição de regressão logística cumulativa (Eq. 1,2) que foram por sua vez utilizadas para calcular a probabilidade de selecionar um genótipo.

$$\hat{\pi}(x_{i1}, x_{i2}, x_{i3}) = \frac{e^{6.6365 - .016NG - 1.114CARVn + 1.472NCn}}{1 + e^{6.6365 - .016NG - 1.114CARVn + 1.472NCn}} \quad Eq 1$$

$$\hat{\pi}(x_{i1}, x_{i2}, x_{i3}, x_{i4}) = \frac{e^{-13.344 + 2.665AC - .146NC - 11.188CARV + .295PMC}}{1 + e^{-13.344 + 2.665AC - .146NC - 11.188CARV + 0.295PMC}} \quad Eq 2$$

A interpretação dos parâmetros foi feita considerando as “odds ratio” (OR) (razões de chance) (HOSMER, LEMESHOW e STURDIVANT, 2013), que é a razão entre a probabilidade de selecionar e não selecionar um clone, calculada a partir da exponencial do coeficiente estimado pela regressão.

$$OR(\hat{\beta}_i) = e^{\hat{\beta}_i}$$

Os modelos foram comparados por meio dos critérios de informação de Akaike (AIC). A Capacidade Preditiva do modelo ou Acurácia (CP): proporção de acertos de

um modelo. Ou seja, é a proporção de verdadeiros-positivos e verdadeiros-negativos em relação ao número de observações, foi calculada da seguinte forma:

$$CP = \frac{VP + VN}{n}$$

Onde VP = verdadeiro positivo; VN = verdadeiro negativo;  $n$  = número de observações.

#### 4.3 RESULTADOS E DISCUSSÃO

Os caracteres qualitativos que apresentaram significância na análise de regressão logística para a seleção massal, aplicada na fase T2, foram NGn, CARVn e NCn (Tabela 4.3 e 4.4). Para o modelo quantitativo, apenas os caracteres AC, NC, CARV e PMC apresentaram efeitos significativos, mostrando que os mesmos têm alguma influência sobre a tomada de decisão do melhorista.

TABELA 4.3 - COEFICIENTES ESTIMADOS PELO MODELO DE REGRESSÃO LOGÍSTICO PARA VARIÁVEIS QUALITATIVAS (MODELO GERAL QUALITATIVO), NA SELEÇÃO MASSAL REALIZADA NA FASE T2 EM CANA ENERGIA.

	Modelo geral qualitativo (1)			
	Estimativa	Erro padrão	Valor z	p-valor
Intercepto	1,577	4,345	0,363	0,716
NG*	-3,798	0,775	-4,902	0,000
CARVn	-1,378	0,604	-2,281	0,022
FER	-0,353	0,298	-1,184	0,236
FAL	-2,235	323,227	-0,007	0,994
CHO	-0,112	0,304	-0,369	0,712
FLO	-0,321	0,301	-1,064	0,287
DES	0,165	0,631	0,261	0,794
DCn	1,201	0,992	1,209	0,226
ACn	1,103	0,661	1,668	0,095
NCn	1,964	0,868	2,262	0,024
PB	0,954	0,849	1,125	0,261
SG	0,449	1,029	0,436	0,663
HAB	-0,338	0,810	-0,417	0,676
AIC		78,445		

\*NG: nota para vigor; CARVn: incidência de carvão; FER: ferrugem marrom; FAL: ferrugem alaranjada; CHO: chochamento; FLO: florescimento; DES: despalha; DCn: diâmetro do colmo; ACn: altura do colmo; NCn: número de colmos; PB: presença de brotação lateral; SG: saliência da gema; HAB: hábito de crescimento e AIC: critérios de *Akaike*.

TABELA 4.4 - COEFICIENTES ESTIMADOS PELO MODELO DE REGRESSÃO LOGÍSTICO PARA VARIÁVEIS QUANTITATIVAS (MODELO GERAL QUANTITATIVO), NA SELEÇÃO MASSAL REALIZADA NA FASE T2 EM CANA ENERGIA.

	Modelo geral quantitativo (2)			
	Estimativa	Erro padrão	Valor z	p-valor
Intercepto	-12,573	3,593	-3,499	0,000
PC*	0,092	0,119	0,772	0,440
FIB	-0,209	0,168	-1,248	0,212
DC	0,159	0,860	0,186	0,853
AC	3,097	0,801	3,867	0,000
NC	0,182	0,047	3,816	0,000
CARV	-0,127	0,042	-3,011	0,002
PMC	0,227	0,100	2,272	0,023
AIC	149,32			

\* PC: teor de sacarose aparente; FIB: teor fibra; DC: diâmetro do colmo; AC: altura do colmo e NC: número de colmos; CARV: incidência de carvão; PMC: peso médio de colmos

Os caracteres FER, FAL, CHO, FLO, DES, DCn, ACn, PB, GEM e HAB para o modelo qualitativos e PC, FIB e DC para o modelo quantitativo não apresentaram coeficientes significativos, sendo descartados do modelo. Os novos modelos (modelos reduzidos) foram feitos utilizando apenas os caracteres que apresentaram significância. De acordo com os valores de razões de chances, os caracteres NCn, CARVn e NG tiveram maior importância para o modelo qualitativo e AC, PMC e NC para o modelo quantitativo (Tabela 4.5).

TABELA 4.5 - PARÂMETROS ESTIMADOS PARA OS MODELOS REDUZIDOS QUALITATIVO E QUANTITATIVO REFERENTE A SELEÇÃO MASSAL REALIZADA NA FASE T2.

	Modelo reduzido qualitativo (3)				
	Estimativa	Erro padrão	valor z	p-valor	OR
Intercepto	7,575	2,016	3,757	0,000	
NG	- 3,982	0,736	-5,409	0,000	0,019
CARVn	-1,052	0,462	-2,279	0,022	0,349
NCn	1,489	0,675	2,204	0,027	4,433
AIC	68,277				
	Modelo reduzido quantitativo (4)				
	Estimativa	Erro padrão	valor z	p-valor	OR
Intercepto	-13,537	2,383	-5,682	0,000	
AC	2,594	0,722	3,589	0,000	13,379
NC	0,170	0,047	3,584	0,000	1,185
CARV	-0,126	0,040	-3,083	0,002	0,881
PMC	0,288	0,076	3,785	0,000	1,335
AIC	145,063				

OR: odds ratio (razões de chance); AIC: critérios de *Akaike*; NG: nota para vigor; CARVn: incidência de carvão; NCn: número de colmos; AC: altura do colmo; NC: número de colmos; CARV: carvão e PMC: peso médio de colmos.

Os coeficientes que apresentam significância na análise de regressão indicam que a variável preditora influencia na decisão de selecionar ou rejeitar uma planta ou

clone. Semelhante à regressão linear múltipla, utiliza-se a seleção de variáveis para eliminar aquelas que não são significativas para o modelo (ZHOU et al., 2014).

Os critérios de informação de “Akaike” (AIC) indicam um melhor ajuste no modelo reduzido para variáveis qualitativas (AIC = 68,277) em comparação com os modelos completos AIC = 78,445 e AIC= 148,710, respectivamente (Tabela 4.3). Os menores valores de AIC refletem um melhor ajuste global (AKAIKE, 1974) (Tabelas 4.3, 4.4 e 4.5). A remoção de características que não contribuem para o melhorista na tomada de decisão contribuiu para um melhor modelo preditivo final.

As plantas que apresentaram nota maior para número de colmos, tem 4,433 vezes mais chance de serem selecionadas para cana energia comparado com uma planta que obteve nota menor para NCn. As plantas que apresentam menor nota para presença de carvão têm 0,349 vezes mais chance de serem selecionadas em relação a uma planta que obteve nota alta (Tabela 4.5).

O número de colmos continua sendo um caractere importante, tendo influência na tomada de decisão na fase T2, especialmente em genótipos destinados à obtenção de biomassa, pois o número de colmos está atrelado a produção de tonelada de colmos por hectare, devido seu alto efeito direto (SILVEIRA et al., 2015). O carvão foi uma doença extremamente importante para a cultura da cana-de-açúcar, visto que pode causar perdas significativas de produtividade na presença da mesma, sendo um caractere importante na tomada de decisão no momento da seleção.

No trabalho realizado por Brasileiro et al. (2016) o carvão foi um caractere que não teve muita importância na seleção dos clones, por tratar-se de outra população, onde possivelmente a incidência da doença foi baixa. No presente trabalho, a incidência de carvão foi maior, fazendo com que a mesma tenha importância para a seleção, sendo utilizada como critério de descarte dos clones.

Em condições de campo, os clones de cana energia apresentaram maior ocorrência das doenças ferrugem marrom e carvão, comparado com clones de cana-de-açúcar (FERNANDES JÚNIOR, 2017). Além disso, o desenvolvimento de cana energia tem ocorrido principalmente no nordeste, onde a incidência de carvão é menor comparada com a região sul onde o trabalho foi desenvolvido.

Na primeira fase de teste (T1), o número de colmos também apresentou maior importância no processo de seleção individual de cana-de-açúcar (ZHOU et al., 2014). Pedrozo et al. (2008) mostraram que plantas que são selecionadas com menos de 6 colmos na fase T1 obtêm baixos rendimentos na fase T2, dessa forma recomendando

a seleção de plantas com mais de 5 colmos. No entanto, outros caracteres, também devem ser levados em consideração para assegurar que os clones selecionados apresentem elevada produtividade de colmos.

A seleção massal realizada a campo selecionou 38 clones, enquanto o modelo qualitativo e o quantitativo selecionaram 42 e 22 clones, respectivamente. Se for considerada a acurácia do modelo, o modelo reduzido qualitativo apresentou capacidade preditiva de 94% e o modelo reduzido quantitativo, capacidade preditiva de 88% (Tabela 4.6).

TABELA 4.6 - CLASSIFICAÇÃO DO NÚMERO DE GENÓTIPOS SELECIONADOS E DESCARTADOS COM A UTILIZAÇÃO DO MODELO DE REGRESSÃO LOGÍSTICO PARA VARIÁVEIS QUALITATIVAS E QUANTITATIVAS.

Fase	Seleção massal	Modelo Reduzido qualitativo (3)			CP
		Seleciona (1)	Descarta (0)	Total	
T2	Seleciona (1)	34	4	38	94%
	Descarta (0)	8	174	182	
	Total	42	178	220	
Fase	Seleção massal	Modelo Reduzido quantitativo (4)			CP
		Seleciona (1)	Descarta (0)	Total	
T2	Seleciona (1)	16	22	38	88%
	Descarta (0)	4	178	182	
	Total	20	220	220	

CP: capacidade preditiva do modelo

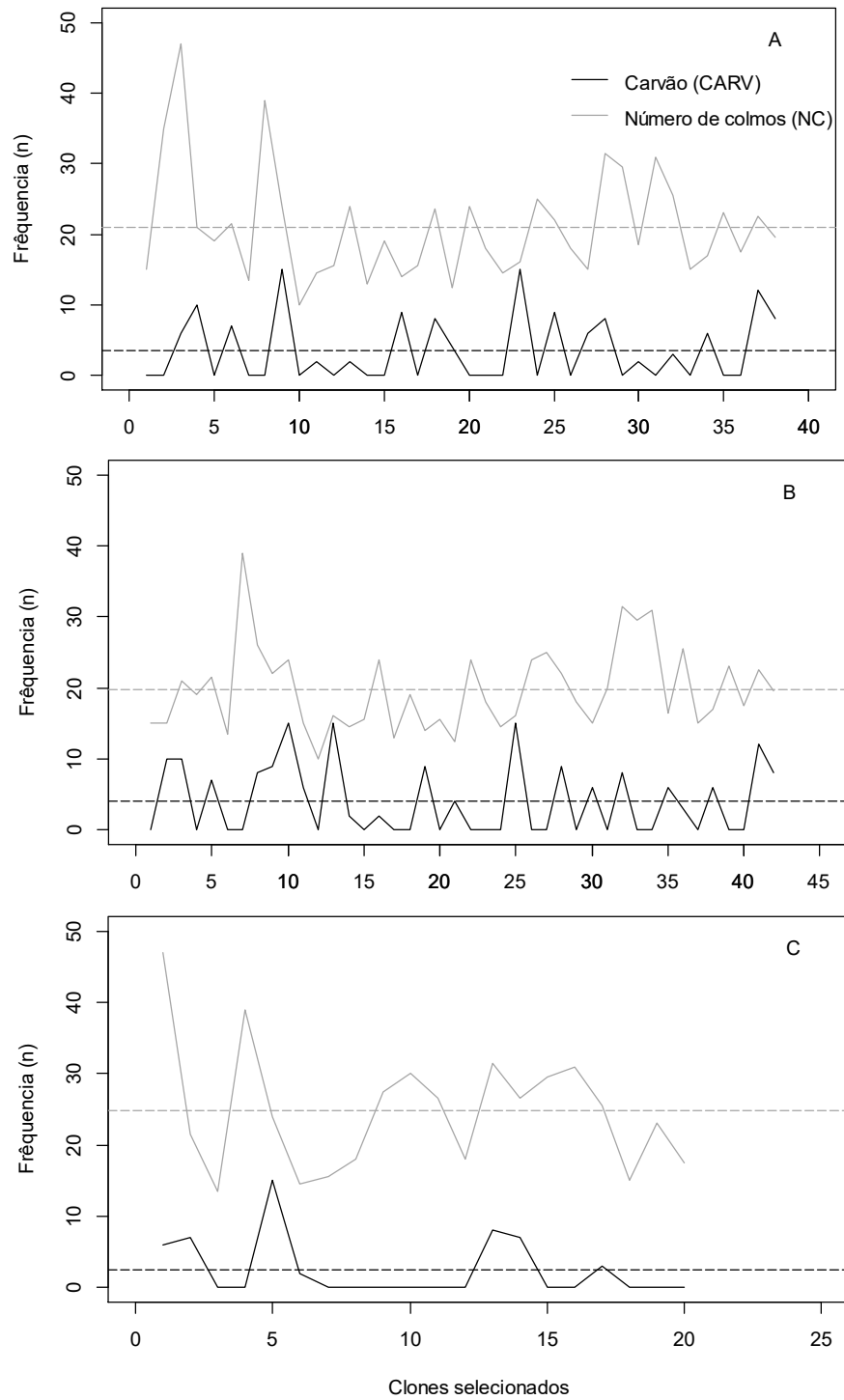
Os clones que foram selecionados pelo método massal, modelo qualitativo e quantitativo apresentaram um máximo de 15 colmos afetados pela doença. A média para número de colmos dos clones selecionados pelo método massal foi de 21,04, para o modelo qualitativo a média foi de 19,90 e para o modelo quantitativo foi 24,7 (Figura 4.1 A, B e C).

Dos 38 clones selecionados a campo pelo método massal, dois deles apresentaram uma incidência de carvão em 15 colmos em relação ao restante dos clones, dos quais, obtiveram baixa incidência (Figura 4.1 A). Foram selecionados clones com incidência de carvão em 15 colmos que obtiveram uma alta produtividade, quanto clones com incidência de carvão em 15 colmos que obtiveram uma menor produtividade.

O modelo qualitativo também selecionou clones com incidência de carvão em 15 colmos, com alta produtividade, quanto clones com incidência de carvão em 15 colmos e baixa produtividade. O modelo quantitativo selecionou um clone com

incidência de carvão em 15 colmos porém, o mesmo apresentou alta produtividade (Figura 4.1 C).

FIGURA 4.1 - NÚMERO DE CLONES SELECIONADOS PELO MÉTODO MASSAL (A), PELO MODELO QUALITATIVO (B) E QUANTITATIVO (C).



Os dois modelos de regressão logística ajustados para a fase T2 em cana energia mostraram-se eficientes na predição dos indivíduos a serem selecionados, pois apresentaram alta acurácia. Porém, o modelo qualitativo selecionou maior número de clones. Apesar dos modelos 3 e 4 apresentarem alta capacidade preditiva, ao analisamos as variáveis levadas em consideração em cada modelo, observamos que o modelo qualitativo pode ser considerado mais prático, pois utiliza variáveis qualitativas (Tabela 4.6).

No melhoramento genético, a utilização de variáveis qualitativas torna o processo mais simples e rápido em relação à utilização de variáveis contínuas, em razão da necessidade de medições, de tempo e de equipamentos. As variáveis explicativas utilizadas para estabelecer o modelo logístico são facilmente e rapidamente avaliadas.

A utilização de caracteres binários ou multicategóricos torna a aplicação de regressão logística interessante, especialmente em programas de melhoramento de cana-de-açúcar para seleção de clones potenciais, levando em consideração a importância de muitos caracteres qualitativos (BRASILEIRO et al., 2016).

O teor de fibra é um caractere importante principalmente quando os programas de melhoramento são voltados para obtenção de materiais de cana energia. Pelo modelo utilizando variáveis qualitativas foram selecionados quatro genótipos a mais em relação à seleção massal (Tabela 4.6). Observa-se que a média para teor de fibra e teor de sacarose manteve-se praticamente a mesma com a utilização do modelo qualitativo (Tabela 4.7).

TABELA 4.7 - MÉDIA DOS GENÓTIPOS SELECIONADOS (1) E NÃO SELECIONADOS (0) PELOS MODELOS QUALITATIVO E QUANTITATIVO E PELA SELEÇÃO MASSAL REALIZADA NA FASE T2 EM CANA ENERGIA.

Seleção	Resposta	PC	FIB
Modelo Qualitativo	0	8,946	14,134
	1	9,281	14,255
Modelo Quantitativo	0	9,013	14,113
	1	8,975	14,556
Seleção Massal	0	8,991	14,151
	1	9,101	14,187

PC: teor de sacarose aparente; FIB: teor de fibra.

Os caracteres que não apresentaram contribuição significativa para a decisão de seleção (Tabela 4.3 e 4.4), também podem receber menor importância durante as decisões de selecionar ou não, otimizando o processo de seleção, além de informar

aos melhoristas quais caracteres possuem menor variabilidade nas populações. Pois de acordo com Zhou et al. (2013) os componentes de produção identificados como não significativos no modelo ajustado, podem indicar os caracteres com menor variabilidade nas populações, podendo ser utilizado para aconselhar estratégias para diversificar caracteres específicos e assim melhorar a variabilidade, bem como a oportunidade de alcançar uma maior seleção e ganhos genéticos.

Os resultados demonstram o potencial da regressão logística em auxiliar no processo de seleção clonal, pois a técnica é capaz de contribuir para a melhor decisão, usando os mesmos critérios utilizados pelos profissionais responsáveis pela seleção durante a segunda fase de teste nos programas de melhoramento genético de cana energia.

#### 4.4 CONCLUSÕES

O número de colmos e a incidência de carvão foram os caracteres de maior importância na tomada de decisão durante a seleção de clones de cana energia na segunda fase de teste. Os dois modelos qualitativo e quantitativo mostraram-se eficientes para prever os genótipos de cana energia a serem selecionados, com acurácia de 94% e 87%, respectivamente, podendo ser utilizados na tomada de decisão durante a seleção de clones.

#### 4.5 REFERÊNCIAS

AGRESTI, A. **An introduction to categorical data analysis**. New York, NY: John Wiley & Sons, 2007.

AKAIKE, H. A New Look at the Statistical Model Identification. **IEEE transactions on automatic control**, v.19, n. 6, 1974.

AMORIM, L.; BERGAMIN FILHO, A.; CARDOSO, C.; MORAES, V.A.; FERNANDES, C.R. Metodologia de avaliação da ferrugem da cana-deaçúcar (*Puccinia melanocephala*). **Boletim Técnico Copersucar**, v.39, p.13-16, 1987.

APARECIDO, L.E.O.; ROLIM, G.S.; RICHETTI, J.; SOUZA, P.S.; JOHANN, J.A. Köppen, Thornthwaite and Camargo climate classifications for climatic zoning in the State of Paraná, Brazil. **Ciência e Agrotecnologia**, v. 40, n.4, p. 405-417, 2016.

BRASILEIRO, B.P.; MARINHO, C.D.; COSTA, P.M.A.; CRUZ, C.D.; PETERNELLI, L.A.; BARBOSA, M.H.P. Selection in sugarcane families with artificial neural networks. **Crop Breeding and Applied Biotechnology**, v.15, p. 72-78, 2015.

BRASILEIRO, B.P.; PETERNELLI, L.A.; SILVEIRA, L.C.I.; BARBOSA, M.H.P. Importance of agronomic traits in the individual selection processo of sugarcane as determined using logistic regression. **Acta Scientiarum Agronomy**, v.38, n.3, p. 289-297, 2016.

BARBOSA, M.H.P.; RESENDE, M. D. V.; BRESSIANI, J. A.; SILVEIRA, L. C. I.; PETERNELLI, L. A. Selection of sugarcane families and parents by Reml/Blup. **Crop Breeding and Applied Biotechnology**, v.5, n.4, p.443-450, 2005.

FERNANDES, A.C. **Cálculos na Agroindústria da cana-de-açúcar**. STAB, 3<sup>a</sup>.ed., 416p, 2011.

FERNANDES JÚNIOR, A.R. **Seleção de famílias e clones de cana energia**. Tese (Doutorado em Agronomia-Produção Vegetal), Universidade Federal do Paraná. Curitiba, 2017.

HOGARTH, D.M.; BERDING, N. Breeding for a better industry: Conventional breeding. **Sugarcane International**, v.24, n., p. 26-31, 2006.

HOSMER, D.W.; LEMESHOW, S.; STURDIVANT, R.X. **Applied logistic regression**. New York: Wiley. 2013.

KIMBENG, C.A.; COX, M.C. Early generation selection of sugarcane families and clones in Australia. **Journal of the American Society of Sugarcane Technologists**, v. 23, p. 20-39, 2003.

KLOSOWSKI, A.C.; RUARO, L.; BESPALHOK FILHO, J.C.; MIO, L.L.M. Proposta e validação de escala para a ferrugem alaranjada da cana-de-açúcar. **Tropical Plant Pathology**, v. 38, n.2, p.166-171, 2013.

MCRAE, T.A.; JACKSON, P.A. Competition effects in selection trials. **Proc Australian Society Sugar Cane Technologists**, v.15, p.177-184, 1998.

PEDROZO, C.A.; BARBOSA, M.H.P.; RESENDE, M.D.V.; PETERNELLI, L.A.; COSTA, P.M.A.; SILVA, F.L. Eficiência da seleção em fases iniciais do melhoramento da cana-de-açúcar. **Revista Ceres**, v.55, n.1, p.1-8, 2008.

R DEVELOPMENT CORE TEAM. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org>, 2016.

SANTOS, H.G.; JACOMINE, P.K.T.; ANJOS, L.H.C.; OLIVEIRA, V.A.; LUBRERAS, J.F.; COELHO, M.R.; ALMEIDA, J.A.; CUNHA, T.J.F.; OLIVEIRA, J.B. **Sistema Brasileiro de Classificação de solos**. 3.ed. Brasília, EMBRAPA, 353p, 2013.

SILVEIRA, L.C.I.; BRASILEIRO, B.P.; KIST, V.; WEBER, H.; DAROS, E.; PETERNELLI, L.A.; BARBOSA, M.H.P. Selection strategy in families of energy cane based on biomass production and quality traits. **Euphytica**, v.204, n.2, p.443-455, 2015.

SKINNER, J.C.; HOGARTH, D.M.; WU, K.K. Selection methods, criteria and indices. In: HEINZ, D. J. (Ed.). **Sugarcane improvement through breeding**. Amsterdam: Elsevier, p. 409-453, 1987.

ZHOU, M.M.; FORESTER, W.; KHOMO, B.; KUPEN, J.; NZIMANDE, L.; NKWANYANA, P.; WALTON, A. Using Logistic Regression Models to aid Selection In Non-Replicated Plots in Sugarcane. **Proc South African Sugarcane Technologists Ass**, v.85, p. 241-245, 2012.

ZHOU, M. Using logistic regression models for selection in non-replicated sugarcane breeding populations. **Euphytica**, v.191, p.415-428, 2013.

ZHOU, M.M.; KIMBENG, C.A.; TEW, T.L.; GRAVOIS, K.A.; PONTIF, M.J.; BISCHOFF, K.P. Logistic regression models to aid selection in early stages of Sugarcane Breeding. **Sugar Tech**, v.16, n.2, p. 150-156, 2014.

## 5 CONSIDERAÇÕES FINAIS

O aumento da demanda por fontes de energia renovável com baixa emissão de carbono, fez com que os programas de melhoramento voltassem os esforços para obtenção de variedades de cana-de-açúcar com maior teor de fibra, denominada cana energia.

No melhoramento genético da cana energia, uma série de critérios são utilizados para selecionar os melhores clones. A baixa precisão experimental, principalmente na primeira e segunda fase de teste, aliada a uma seleção massal baseada no fenótipo, instiga a busca de estratégias que auxiliem no momento da seleção.

A utilização de modelos matemáticos com objetivo de auxiliar os programas de melhoramento genético na seleção de clones tem se mostrado eficiente, podendo ser utilizado para auxiliar na seleção de clones de cana energia. A atribuição de notas para os caracteres possibilita otimizar o processo de seleção, maximizando o tempo e os recursos utilizados nas coletas de dados nas fases a campo.

Em trabalhos futuros é necessário testar a aplicabilidade das técnicas avaliadas visando selecionar clones que atendam às necessidades do setor sucroenergético.

## 6 REFERÊNCIAS GERAIS

- AFIFI, A.; MAY, S.; CLARK, V.A. **Practical Multivariate Analysis**. 5ª ed, 2012.
- AGRESTI, A. **An introduction to categorical data analysis**. 2ª ed. Hoboken: John Wiley & Sons, 2007.
- AKAIKE, H.A New Look at the Statistical Model Identification. **IEEE transactions on automatic control**, v.19, n. 6, 1974.
- AMORIM, L.; BERGAMIN FILHO, A.; CARDOSO, C.; MORAES, V.A.; FERNANDES, C.R. Metodologia de avaliação da ferrugem da cana-deaçúcar (*Puccinia melanocephala*). **Boletim Técnico Copersucar**, v.39, p.13-16, 1987.
- APARECIDO, L.E.O.; ROLIM, G.S.; RICHETTI, J.; SOUZA, P.S.; JOHANN, J.A. Köppen, Thornthwaite and Camargo climate classifications for climatic zoning in the State of Paraná, Brazil. **Ciência e Agrotecnologia**, v. 40, n.4, p. 405-417, 2016.
- AUSTIN, J.T.; YAFFEE, R.A.; HINKLE, D.E. Logistic regression for research in higher education. **Higher Education: Handbook of Theory and Research**, v.8, p.379-410, 1992.
- BAFFA, D.C.; COSTA, P.M.A.; SILVEIRA, G.; LOPES, F.J.F.; BARBOSA, M.H.P.; LOUREIRO, M.E.; CRUZ, C.D.; PETERNELLI, L.A. Path analysis for selection of Saccharification-Efficient sugarcane genotypes through agronomic traits. **Agronomy Journal**, v. 106, n.5, 2014.
- BARBOSA, M.H.P.; RESENDE, M.D.V.; BRESSIANI, J.A.; SILVEIRA, L.C.I.; PETERNELLI, L.A. Selection of sugarcane families and parents by Reml/Blup. **Crop Breeding and Applied Biotechnology**, v.5, n.4, p.443-450, 2005.
- BARBOSA, M.H.P.; SILVEIRA, L.C.I. Breeding and cultivar recommendations. In Santos F, Borém A and Caldas C (eds.) **Sugarcane: bioenergy, sugar and ethanol: technology and prospects**. MAPA/ACS: UFV/DEA, Brasília, p. 301-318, 2012.
- BELLO, O.B., S.Y. ABDULMALIQ, M.S. AFOLABI.; S.A. IGE. Correlation and path coefficient analysis of yield and agronomic characters among open pollinated maize varieties and their F1 hybrids in a diallel cross. **African Journal Biotechnology**. v.9, p.2633-2639, 2010.
- BRASILEIRO, B.P.; PETERNELLI, L.A.; BARBOSA, M.H.P. Consistency of the results of path analysis among sugarcane experiments. **Crop Breeding and Applied Biotechnology**, v. 13, p.99-107, 2013.
- BRASILEIRO, B.P.; MARINHO, C.D.; COSTA, P.M. A.; CRUZ, C.D.; PETERNELLI, L.A.; BARBOSA, C.H.P. Selection in sugarcane families with artificial neural networks. **Crop Breeding and Applied Biotechnology** v.15, p.72-78, 2015.

BRASILEIRO, B.P.; PETERNELLI, L.A.; SILVEIRA, L.C.I.; BARBOSA, M.H.P. Importance of agronomic traits in the individual selection processo of sugarcane as determined using logistic regression. **Acta Scientiarum Agronomy**, v.38, n.3, p. 289-297, 2016.

BORÉM, A. **Melhoramento de espécies cultivadas**. Viçosa: UFV 2 ed, 969p, 2005.

CABRERA, A.F. Logistic regression analysis in higher education: An applied perspective. **Higher Education: Handbook of Theory and Research**, Vol. 10, 225-256, 1994.

CARVALHO-NETTO, O.V.; BRESSIANI, J.A.; SORIANO, H.L.; FIORI, C.S.; SANTOS, J.M.; BARBOSA, G.V.S.; XAVIER, M.A.; LANDELL, M.G.A.; PEREIRA, G.A.G. The potential of the energy cane as the main biomass crop for the cellulosic industry. **Chemical and Biological Technologies in Agriculture**, v.1, p.1-8, 2014.

CASELLA, G.; BERGER, R. **Interferência estatística**. 2<sup>nd</sup> ed. São Paulo: C Learning, 588p, 2001.

CESNIK, R.; MIOCQUE, J. **Melhoramento da cana-de-açúcar**. Embrapa Informação Tecnológica, 307.p, 2004.

CHARNET, R. **Análise de modelos de regressão linear**. Campinas: Unicamp, 357p, 2008.

CHAUDHARY, R.R.; JOSHI, B.K. Correlation and Path Coefficient Analyses in Sugarcane. **Nepal Agriculture Research Journal**, v. 6, 2005.

COSTA NETO, P.L.O. **Estatística**. 2. Ed. São Paulo: E. Blucher, 280p, 2009.

COX, D.R.; SNELL, E.J. **The analysis of binary data** (2nd ed.). London: Chapman and Hall, 1989.

CRESTE, S.; ROSA JUNIOR, V.E.; PINTO, L.R.; ALBINO, J.C.; FIGUEIRA, A.V.O. A Biotecnologia como ferramenta para o melhoramento genético. In: DINARDO MIRANDA, L.L.; VASCONCELOS, A.C.M.; LANDELL, M.G.A. **Cana-de-açúcar**. Instituto agrônômico. p.156-176. 2008.

CRUZ, C.D. Programa GENES: **análise multivariada e simulação**. Viçosa: UFV, 175p, 2006.

CRUZ, C.D.; CARNEIRO, P.C.S.; REGAZZI, A.J. **Modelos biométricos aplicados ao melhoramento genético**. 3<sup>a</sup> ed. UFV, Vicosa, MG, Brazil, 2014.

DAL-BIANCO, M.; CARNEIRO, M.S.; HOTTA, C.T.; CHAPOLA, R.G.; HOFFMANN, H.P.; GARCIA, A.A.F.; SOUZA, G.M. Sugarcane improvement: how far can we go?. **Elsevier Current Opinion in Biotechnology**, v.23, p.265-270. 2012.

DINARDO-MIRANDA, L.L.; VASCONCELOS, A.C.M.; LANDELL, M.G.A. **Cana-de-açúcar**. Campinas: Instituto Agrônômico, 882.p, 2008.

ENSINAS, A.V.; NEBRA, S.A.; LOZANO, M.A.; SERRA, L.M. Analysis of process steam demand reduction and electricity generation in sugar and ethanol production from sugarcane. **Energy Conversion and Management**, v.48, p.2978-2987, 2007.

ESPÓSITO, D.P.; PETERNELLI, L.A.; DE PAULA, T.O.M.; BARBOSA, M.H.P. Path analysis using phenotypic and genotypic values for yield components in the selection of sugarcane families. **Ciência Rural**, v.42, n.1, p.38-44, 2012.

FALCONER, D.S. **Introdução a genética quantitativa**. Viçosa, UFV, Impr.Univ, 279p, 1987.

FALCONER, D.S.; MACKAY, T.F.C. **Introduction to qualitative genetics**. 4.ed. New York, Longman. p.464, 1996.

FERNANDES, A.C. **Cálculos na Agroindústria da cana-de-açúcar**. STAB, 3ª.ed., 416p, 2011.

FERNANDES JÚNIOR, A.R. **Seleção de famílias e clones de cana energia**. Tese (Doutorado em Agronomia-Produção Vegetal), Universidade Federal do Paraná. Curitiba, 2017.

FERREIRA, F.M.; BARROS, W.S.; SILVA, F.L.; BARBOSA, M.H.P.; CRUZ, C.D.; BASTOS, I.T. Relação fenotípica e genotípica entre componentes de produção em cana-de-açúcar. **Bragantia**, v.66, n.4, p.605-610, 2007.

HAIR, J.F.; BLACK, W.C.; BABIN, B.J.; ANDERSON, R.E.; TATHAM, R.L. **Análise Multivariada de Dados**. Tradução Adonai Schlup Sant'Anna. -6. Ed.- Porto Alegre, 688.p, 2009.

HEATON, E.A., DOHLEMAN, F.G. AND LONG, S.P. Meeting US biofuel goals with less land: the potential of Miscanthus. **Global Chang Biology**, v.14, p.2000-2014, 2008.

HOGARTH, D.M.; BERDING, N. Breeding for a better industry: Conventional breeding. **Sugarcane International**, v.24, n., p. 26-31, 2006.

HOSMER, D.W.; LEMESHOW, S.; STURDIVANT, R.X. **Applied logistic regression**. New York: Wiley. 2013.

JAMES, N.I. Yield components in random and selected sugarcane populations. **Crop Science**, v.11, p.906-908, 1971.

LANDELL, M.G.A.; BRESSIANI, J. A. **Melhoramento genético, caracterização e manejo varietal**. In: DINARDO-MIRANDA, L.L.; VASCONCELOS, A.C. M.; LANDELL, M. G. A. Cana-de-açúcar. Campinas, SP: Instituto Agrônômico. p.101-155. 2005.

KENNEDY, A.J. Prospects for combining high sucrose content with increased fibre to generate multi-purpose cane varieties. In: **Proceeding of the Conference of West Indies Sugar Technologists**, Jamaica, 2008.

KIM, M.; DAY, D.F. Composition of sugar cane, energy cane, and sweet sorghum suitable for ethanol production at Louisiana sugar mills. **Journal of Industrial Microbiology and Biotechnology**, v. 38, n. 7, p. 803-807, 2011.

KIMBENG, C.A.; COX, M.C. Early generation selection of sugarcane families and clones in Australia. **Journal of the American Society of Sugarcane Technologists**, v. 23, p. 20-39, 2003.

KLOSOWSKI, A.C.; RUARO, L.; BESPALHOK FILHO, J.C.; MIO, L.L.M. Proposta e validação de escala para a ferrugem alaranjada da cana-de-açúcar. **Tropical Plant Pathology**, v. 38, n.2, p.166-171, 2013.

LEI, P.W.; KOEHLI, L.M. **Linear discriminant analysis versus logistic regression: A comparison of classification errors**. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA, 2000.

LI, C.C. Concept of path coefficient and its impact on population genetics. **Biometrics**, v.12, p.190-210, 1975.

LOUREIRO, M.E.; BARBOSA, M.H.P.; LOPES, F.J.F.; SILVÉRIO, F.O. Sugarcane breeding and selection for more efficient biomass conversion in cellulosic ethanol. In: Buckeridge MS, Goldman GH (eds) **Routes to cellulosic ethanol**. Springer, New York, pp 199-239, 2011.

MATSUOKA, S.; BRESSIANI, J.; MACCHERONI, W.; FOUTO, I. Bioenergia de cana. In: SANTOS, F.; BORÉM, A.; CALDAS, C. (Org). **Cana-de-açúcar: Bioenergia, Açúcar e Álcool**. 1ed. Viçosa: UFV, v1. p:547-577, 2010.

MARCHEZAN, E.; MARTIN, T.N.; SANTOS, F.M. DOS.; CAMARGO, E.R. Análise de coeficiente de trilha para os componentes de produção em arroz. **Ciência Rural**, v.35, n.5, p.1027-1033, 2005.

MCRAE, T.A.; JACKSON, P.A. Competition effects in selection trials. **Proc Australian Society Sugar Cane Technologists**, v.15, p.177-184, 1998.

MILLIGAN, S.B.; GRAVOIS, K.A.; BISCHOFF, K.P.; MARTIN, F.A. Crops effects on broad-sense heritabilities and genetic variances of sugarcane yield components. **Crop Science**. v.30, p.344-349, 1990.

MONTGOMERY, D.C.; PECK, E.A.; VINING, G.G. **Introduction to linear regression analyses**. Wiley, p. 504, 2012.

NARDINO, M.; SOUZA, V.Q.; BARETTA, V.A. KONFLANZ.; CARVALHO, I.R.; FOLLMANN, D.N.; CARON, B.O. Association of secondary traits with yield in maize F1's. **Ciência Rural** v.46, p.776-782, 2016.

KUTNER, M.H.; NACHTSHEIM, J.C.; NETER, J.; LI, W. **Applied linear statistical models**. 5<sup>th</sup> ed. New York: McGraw- Hill/Irwin, 1396 p, 2005.

OLIVEIRA, R.A.; DAROS, E.; RESENDE, M.D.V.; BESPALHOK-FILHO, J.C.; ZAMBON, J.L.C.; SOUZA, T.R.; LUCIUS, A.S.F. Procedimento Blupis e seleção massal em cana-de-açúcar. **Bragantia**, v. 70, n. 4, p.796-800, 2011.

OLIVOTO, T.; SOUZA, V.Q de.; NARDINO, M.; CARVALHO, I.R.; FERRARI, M.; PELEGRIN, A.J de.; SZARESKI, V.J.; SCHMIDT, D. Multicollinearity in Path Analysis: A Simple Method to Reduce Its Effects. **Agronomy Journal**, v.1 0 9, n.1, 2017.

PEDROZO, C.A.; BARBOSA, M.H.P.; RESENDE, M.D.V.; PETERNELLI, L.A.; COSTA, P.M.A.; SILVA, F.L. Eficiência da seleção em fases iniciais do melhoramento da cana-de-açúcar. **Revista Ceres**, v.55, n.1, p.1-8, 2008.

PENG, C.Y.J.; LEE, K.L.; INGERSOLL, G.M. An introduction to logistic regression analysis and reporting. **The Journal of Educational Research**, v.96, n.1, p. 3-14, 2013.

PENG, C.Y.; SO, T.S.H. Modeling strategies in logistic regression. **Journal of Modern Applied Statistical Methodr**, v.14, p.147-156, 2002.

PENG, C.Y.; SO, T.S.; STAGE, F.K.; ST. JOHN, E.P. The use and interpretation of logistic regression in higher education journals: 1988-1999. **Research in Higher Education**, v.43, p.259-293, 2002.

PRESS, S.J.; WILSON, S. Choosing between logistic regression and discriminant analysis. **Journal of the American Statistical Association**. 73, 699-705, 1978.

R DEVELOPMENT CORE TEAM. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org>, 2016.

RAMOS, R.S.; BRASILEIRO, B.P; SILVEIRA, L.C.I.; KIST, V.; PETERNELLI, L.A.; BARBOSA, M.H.P. Selection of parents, families and clones to obtain energy cane. **Agronomy Journal**, v.109, p.1-7, 2017.

REDDY, C.R.; REDDI, M.V. Degree of genetic determination, correlation and genotypic and phenotypic path analysis of cane and sugar yield in sugarcane. **Indian Journal of Genetics and Plant Breeding**, v.46, p.550-557, 1986.

RESENDE, M.D.V. Software Selegen-REML/BLUP: a useful tool for plant breeding. **Crop Breeding and Applied Biotechnology** v.16, p.330 -339, 2016.

RESENDE, M.D.V.; BARBOSA, M.H.P. Selection via simulated BLUP based on family genotypic effects in sugarcane. **Pesquisa Agropecuária Brasileira**, v.41, p.421-429, 2006.

RESENDE, M.D.V. **SELEGEN-REML/BLUP Sistema Estatístico e seleção genética Computadorizada Via Modelos Lineares Mistos**. Ed Embrapa, 1ª ed, 360p, 2007.

RESENDE, M.D.V. Genética biométrica e estatística no melhoramento de plantas perenes. **Embrapa Informação Tecnológica**: Brasília, 975p, 2002.

SANTCHURN, D.; RAMDOYAL, K.; BADALOO, M.G.H.; LABUSCHAGNE, M. From sugar industry to cane industry: investigations on multivariate data analysis techniques in the identification of different high biomass sugarcane varieties. **Euphytica**, v.185, p.543-558, 2012.

SANTOS, H.G.; JACOMINE, P.K.T.; ANJOS, L.H.C.; OLIVEIRA, V.A.; LUBRERAS, J.F.; COELHO, M.R.; ALMEIDA, J.A.; CUNHA, T.J.F.; OLIVEIRA, J.B. **Sistema Brasileiro de Classificação de solos**. 3.ed. Brasília, EMBRAPA, 353p, 2013.

SILVA, F.L.; PEDROZO, C.A.; BARBOSA, M.H.P.; RESENDE, M.D.V.; PETERNELLI, L.A.; COSTA, P.M DE A.; VIEIRA, M.S. Análise de trilha para os componentes de produção de cana-de-açúcar via blup. **Revista Ceres**, v.56, n.3, p.308-314, 2009.

SILVEIRA, L.C.I.; BRASILEIRO, B.P.; KIST, V.; WEBER, H.; DAROS, E.; PETERNELLI, L.A.; BARBOSA, M.H.P. Selection strategy in families of energy cane based on biomass production and quality traits. **Euphytica**, v.204, n.2, p.443-455, 2015a.

SILVEIRA, L.C.I.; BRASILEIRO, B.P.; KIST, V.; DAROS, E.; PETERNELLI, L.A. Genetic diversity and coefficient of kinship among potential genitors for obtaining cultivars of energy cane. **Revista Ciência Agronômica**, v. 46, n. 2, p. 358-368, 2015b.

SIEGEL, S.; CASTELLAN JR, N.J. **Estatística não-paramétrica: para ciência do comportamento**. Porto Alegre: Artmed 2ª Ed, 448p, 2006.

SKINNER, J.C.; HOGARTH, D.M.; WU, K.K. Selection methods, criteria and indices. In: HEINZ, D. J. (Ed.). **Sugarcane improvement through breeding**. Amsterdam: Elsevier, p. 409-453, 1987.

SUKHCHAIN, S.D.; NSAINI, G.S. Inter-relationships among cane yield and commercial cane sugar and their component traits in autumn plant crop of sugarcane. **Euphytica**, v.95, p. 109-114, 1997.

TABACHNICK, B.G.; FIDELL, L.S. **Using multivariate statistics** (4ª ed.). Needham Heights, MA: Allyn & Bacon, 2001.

TENA, E.; MEKBIB, F.; AYANA, A. Correlation and Path Coefficient Analyses in Sugarcane Genotypes of Ethiopia. **American Journal of Plant Sciences**, v.7, n.10, p.1490-1497, 2016.

TOLLEFSON, J. Energy: not your father's biofuels. **Nature**, v.451, p.880-883, 2008.

TYAGI, A.P.; LAL, P. Correlation and path coefficient analysis in sugarcane. **The South Pacific Journal of Natural Science**, v.25, p.1-9, 2007.

ZENI-NETO, H, DAROS, E, BESPALHOK-FILHO J.C, SCAPIM C.A, VIDIGAL, M.C.G, FILHO, P.S.V. Selection of families and parents of sugarcane (*Saccharum* spp.) through mixed models by joint analysis of two harvests. **Euphytica**, v.193, p.391-408, 2013.

ZHOU, M.M.; FORESTER, W.; KHOMO, B.; KUPEN, J.; NZIMANDE, L.; NKWANYANA, P.; WALTON, A. Using Logistic Regression Models to aid Selection In Non-Replicated Plots in Sugarcane. **Proc South African Sugarcane Technologists Ass**, v.85, p. 241-245, 2012.

ZHOU, M. Using logistic regression models for selection in non-replicated sugarcane breeding populations. **Euphytica**, v.191, p.415-428, 2013.

ZHOU, M.M.; KIMBENG, C.A.; TEW, T. L.; GRAVOIS, K.A.; PONTIF, M.J.; BISCHOFF, K. P. Logistic regression models to aid selection in early stages of Sugarcane Breeding. **Sugar Tech**, v.16, n.2, p. 150-156, 2014.

WACLAWOVSKY, A.J.; SATO, P.M.; LEMBKE, C.G.; MOORE, P.H.; SOUZA, G.M. Sugarcane for bioenergy production: an assessment of yield and regulation of sucrose content. **Plant Biotechnology Journal**, v.8, p. 263-276, 2010.

WRIGHT, S. Correlation and causation. **Journal of Agricultural Research**, v.20, n.7, p.557-585, 1921.

## APÊNDICE 1- Comandos análise de trilha

---

```

matriz.de.efeitos.diretos<-function(matriz.de.correlacao,numero.de.variaveis)
{ m.cor<-matriz.de.correlacao
  n<-numero.de.variaveis
  r<-m.cor[2:n,2:n]
  a<-solve(r)
  b<-m.cor[1,2:n]
  solucao<-a%*%b
  return(solucao)}

matriz.de.correlacao.variaveis.explicativas<-function(matriz.de.correlacao,numero.de.variaveis)
{ m.cor<-matriz.de.correlacao
  n<-numero.de.variaveis
  r<-m.cor[2:n,2:n]
  return(r)}

m.cor<-cor(dados)
n<-6
m.cor
matriz.de.efeitos.diretos(m.cor,n)
r<-(matriz.de.correlacao.variaveis.explicativas(m.cor,n))
determinante<-det(r) # determinate da matriz
determinante

e<-eigen(r)
Nc<-max(e$values)/min(e$values)
Nc

x<-matriz.de.efeitos.diretos(m.cor,n)
x
y<-cbind(x,x,x,x,x)
y
y*r

R2<-((x[1]*m.cor[1,2])+(x[2]*m.cor[1,3])+(x[3]*m.cor[1,4])+(x[4]*m.cor[1,5])+(x[5]*m.cor[1,6]))
R2

efeito.residual<-sqrt(1-R2)
efeito.residual

```

## APÊNDICE 2 - Comandos análise de regressão logística

---

```
mod1<-glm(SEL ~ NG+CARV+FER+FAL+CHO+FLO+DES+DC+AP+NC+CT+GEM+HAB, data =  
dados, family = binomial(link="logit"))  
summary(mod1)
```

```
mod2<-glm(SEL ~NG+CARV+NC, data = dados, family = binomial(link="logit"))  
summary(mod2)  
exp(coef(mod2))
```

---

```
xtabs(~SEL+PROB, data = dados1)  
CP<-((172+34)/(172+34+4+8))  
CP
```