



Simposio de Integração Acadêmica



Universidade Federal de Viçosa

Programas GENES e R: Análise de Experimentos



Vinícius Quintão Carneiro
vqcarneiro@gmail.com
www.fenomica.com.br



APLICATIVO GENES

estatística Experimental

Biometria

Análise Multivariada

Diversidade Genética

Simulação

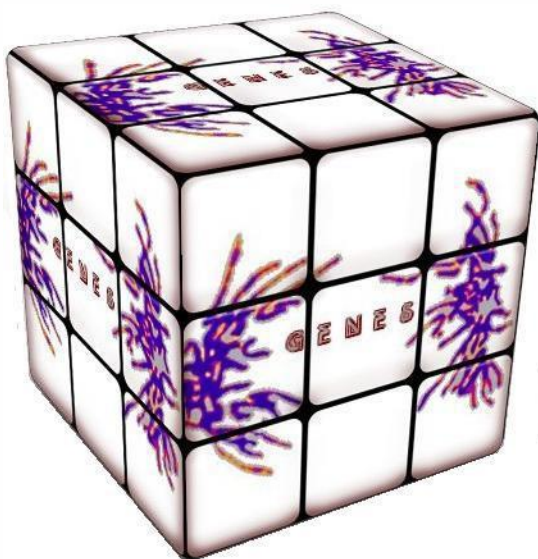
Matrizes

Integração

Inteligência Computacional

Fenômica

Genômica



Calculadora

Estatística

Correlação

Matriz

GenesCalc

Gráfico



Atualização



Autor



Citação

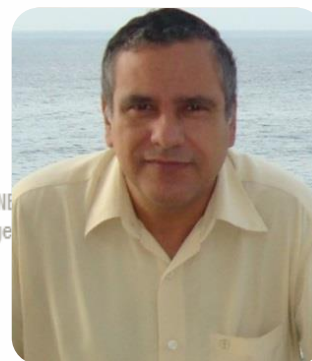
Cruz, C.D. GENES: um aplicativo para análise de dados em estatística experimental e genética quantitativa. Genética 51(2):271-276, 2013



Rede social



Tutoriais



Universidade



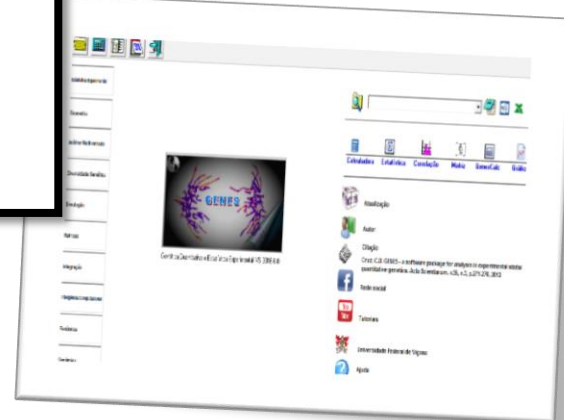
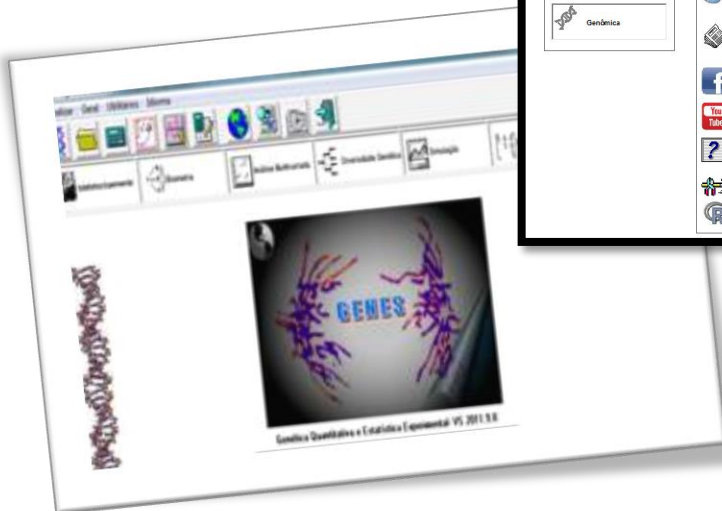
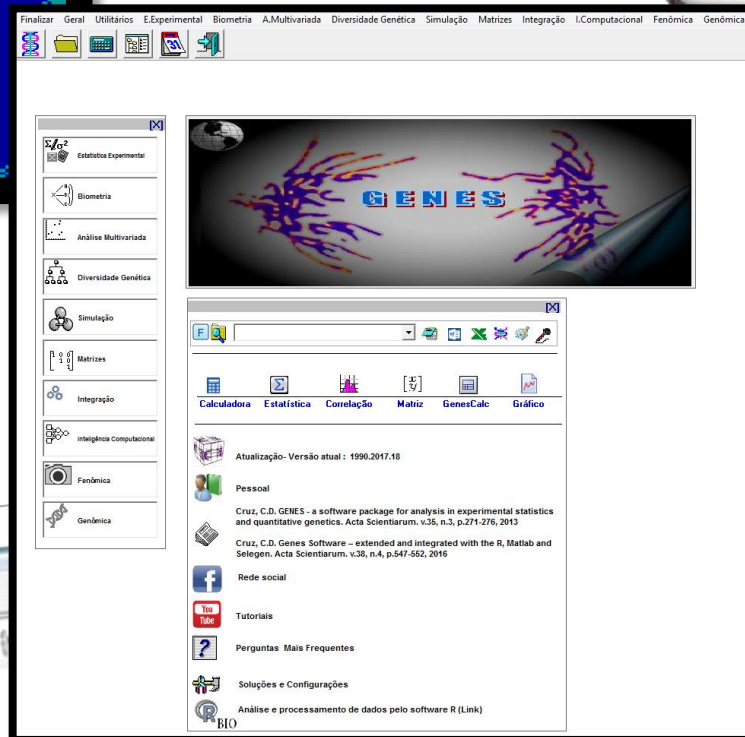
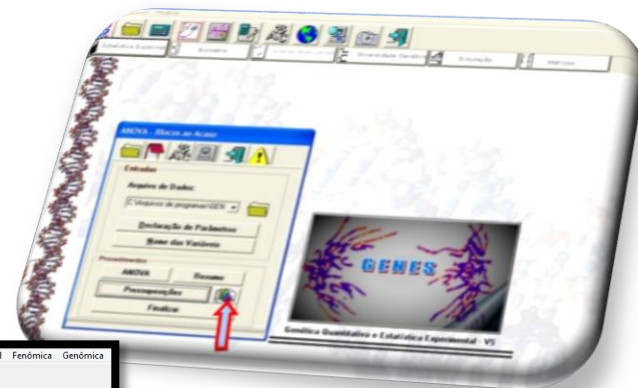
Universidade



GBOL



GPOP



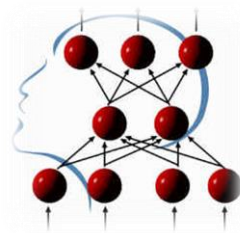
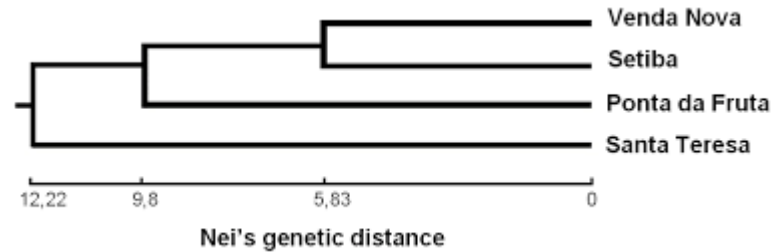
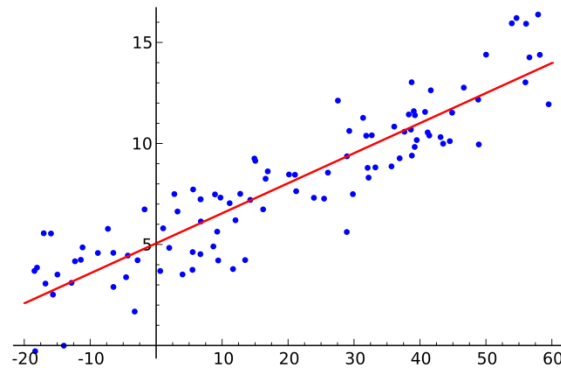
Rotinas



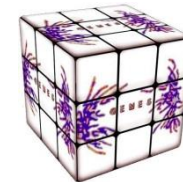
- Geral
- Utilitários
- Estatística Experimental
- Biometria
- Análise Multivariada
- Diversidade Genética
- Simulação
- Matrizes
- Inteligência Computacional
- Fenômica
- Genômica



GQMOL



Ajuda



Facebook interface for Genes News.

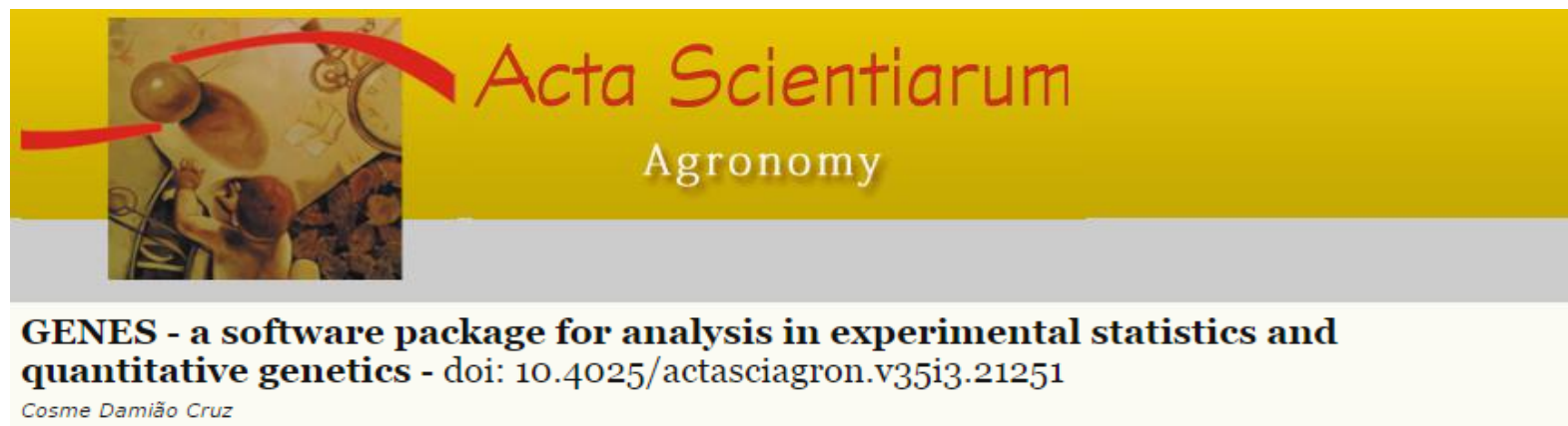
Top navigation bar: f Genes News [Search icon] Vinicius Página inicial

Secondary navigation bar: Página Mensagens Notificações 1 Informações Ferramentas de publicação Co

Main content area:

- Profile picture: A 3D cube with a grid pattern and purple floral designs.
- Cover photo: Four book covers for the "Programa GENES" series by Cosme Damiano Cruz:
 - Programa GENES ESTATÍSTICA EXPERIMENTAL E MATRIZES
 - Programa GENES ANÁLISE MULTIVARIADA E SIMULAÇÃO
 - Programa GENES BIOMETRIA
 - Programa GENES DIVERSIDADE GENÉTICA
- Page name: Genes News @GenesNews
- Interaction buttons: Enviar mensagem (circled in red), Curtiu, Mensagem, ...
- Bottom navigation bar: Linha do Tempo (underlined), Sobre (circled in red), Fotos, Eventos, Mais, + Adicionar seção Loja

<https://www.youtube.com/user/roldoagro>



Acta Scientiarum



<http://www.uem.br/acta>
ISSN printed: 1679-9275
ISSN on-line: 1807-8621
Doi: 10.4025/actasciagron.v38i4.32629

Genes Software – extended and integrated with the R, Matlab and Selegen

Cosme Damião Cruz



1. CRUZ, C. D. . Programa Genes - Análise multivariada e simulação. 1. ed. Viçosa, MG: Editora UFV, 2006. v. 1. 175 p.
2. CRUZ, C. D. . Programa Genes - Biometria. 1. ed. Viçosa, MG: Editora UFV, 2006. v. 1. 382 p.
3. CRUZ, C. D. . Programa Genes - Diversidade Genética. 1. ed. Viçosa, MG: Editora UFV, 2008. v. 1. 278 p.
4. CRUZ, C. D. . Programa Genes - Estatística Experimental e Matrizes. 1. ed. Viçosa: Editora UFV, 2006. v. 1. 285 p.

Software



ca

Arquivo Editar Pacotes Janelas Ajuda



```
R Console

R version 3.4.0 (2017-04-21) -- "You Stupid Darkness"
Copyright (C) 2017 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R é um software livre e vem sem GARANTIA ALGUMA.
Você pode redistribuí-lo sob certas circunstâncias.
Digite 'license()' ou 'licence()' para detalhes de distribuição.

R é um projeto colaborativo com muitos contribuidores.
Digite 'contributors()' para obter mais informações e
'citation()' para saber como citar o R ou pacotes do R em publicações.

Digite 'demo()' para demonstrações, 'help()' para o sistema on-line de ajuda,
ou 'help.start()' para abrir o sistema de ajuda em HTML no seu navegador.
Digite 'q()' para sair do R.

> |
```

```
Sem nome - Editor R

train.net<-function(n_training,amostra,rang,decay, maxit,erro )
{
  i<-1
  size1<-seq(1,15)
  size2<-matrix(0,10000)
  erro<-matrix(1,10000)
  repeat
  {
    a<-data.frame(grupos[,-1],g = factor(grupos[,1]))
    samp<-c(sample(1:n_training,amostra), sample((n_training+1):(2*n_training),amos
    ncam<-sample(size1,1)
    size2[i,1]<-ncam
    redel<- nnet(g~ ., data = a, subset = samp, size=ncam, rang =rangg,decay = dec
    b<-table(a$g[-samp], predict(redel, a[-samp,], type = "class"))
    erro[i]<-1-(sum(diag(b))/600)
    i<-i+1
    if(erro[i-1]<erroo)
      break
  }
}
train.net(n_training,amostra,rang,decay, maxit,erro)
#####
```



The screenshot shows the RStudio interface with the following components highlighted:

- 1- Code Editor:** The top-left pane showing R code for loading ggplot2, viewing the diamonds dataset, and creating a plot.
- 2- R Console:** The bottom-left pane showing the output of the code, including summary statistics for the diamonds dataset.
- 3- Workspace and History:** The top-right pane showing the current workspace with the 'diamonds' dataset and its summary.
- 4- Plots and files:** The bottom-right pane showing a scatter plot titled 'Diamond Pricing' with 'Carat' on the x-axis and 'Price' on the y-axis, colored by clarity.

```
1 library(ggplot2)
2
3 view(diamonds)
4 summary(diamonds)
5
6 summary(diamonds$price)
7 aveSize <- round(mean(diamonds$carat), 4)
8 cla
9
10 p <-
11
12
13 main="diamond Pricing")
14
```

Console

	x	y	z
Min.	0.000	0.000	0.000
1st Qu.	4.710	4.720	2.910
Median	5.700	5.710	3.530
Mean			3.539
3rd Q			4.040
Max.			1.800

Diamond Pricing

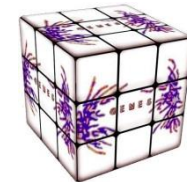
Price

Carat

Integração



Roteiro do Curso



1. Instalação, Configuração e Procedimentos Básicos

❖ Integração

2. Estatística Experimental

- Estatística Experimental
- Planilha de Experimentos
- ANOVA – DIC / DBC
- Testes de comparação entre médias
- Correlações

3. Fitotectnia

- Regressão
- Fatorial

4. Melhoramento Vegetal

- Delineamentos de Blocos Incompletos
 - Látice
 - Blocos Aumentados
- Análise de grupos de experimentos

Instalação



- ❖ Download: Genes News: <https://www.facebook.com/GenesNews/>
Site: <ftp://ftp.ufv.br/dbg/biodata/>

- Arquivos:

1. bioinfo_setup -> Instalação propriamente dita
2. genesexe -> Descompactar em C:
3. dados -> Descompactar em C:

Atenção: Só irá funcionar se fizer estes três procedimentos

INSTALAÇÃO



bioinfo_setup

INSTALAÇÃO/ATUALIZAÇÃO



genesexe



dados

Roteiro - Instalação

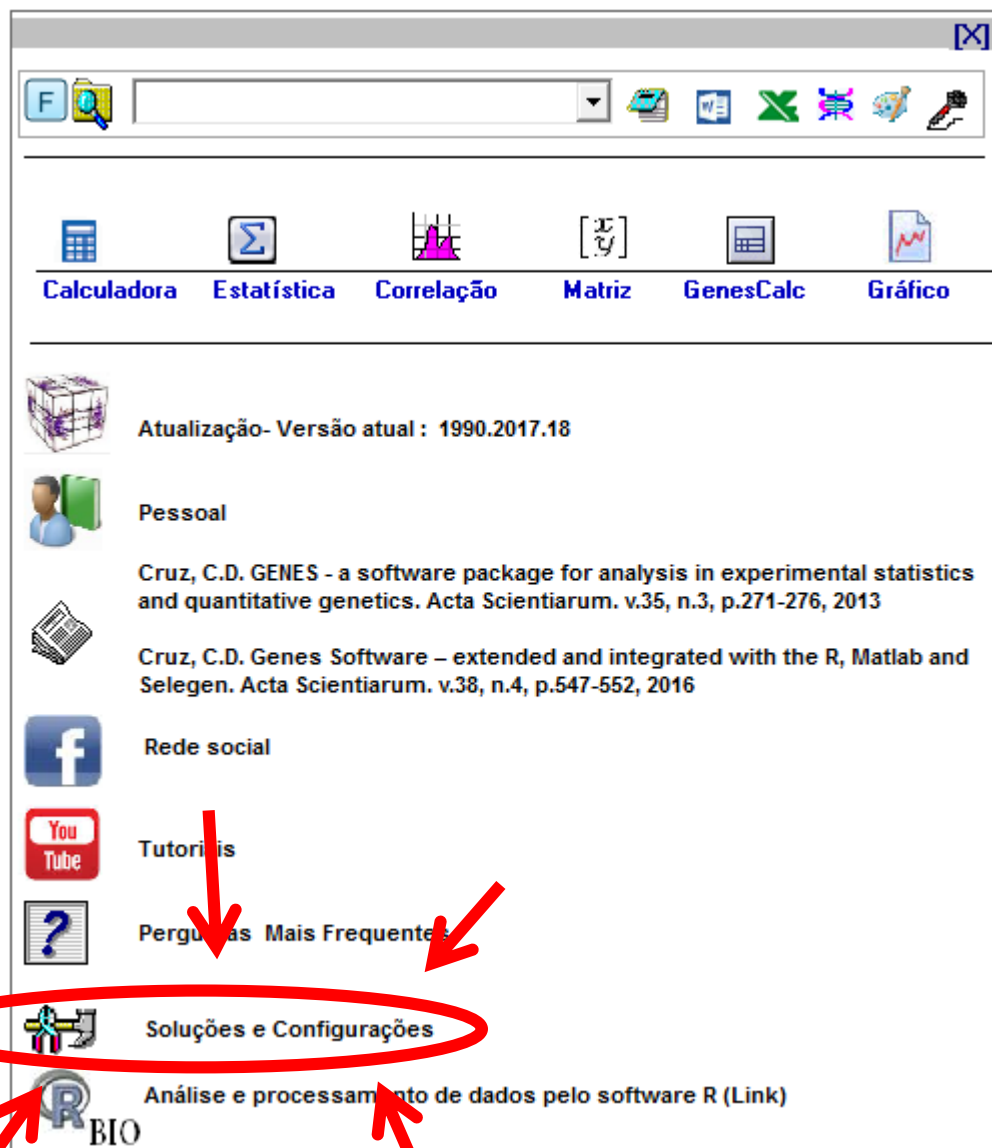


1. Se você vai utilizar o programa genes pela primeira vez, recomendamos instalar um programa auxiliar denominado bioinfo. Faça o seguinte:
 - a. Faça download do arquivo bioinfo_setup.zip
 - b. Descompacte a pasta
 - c. Execute o programa bioinfo_setup

2. Se você está fazendo atualização, faça o seguinte:
 - a. Faça download do arquivo genesexe.zip
 - b. Descompacte a pasta em c:\genesexe. Posteriormente, encontre, dentro da pasta, o arquivo genes.exe e execute-o. Crie um atalho na área de trabalho para facilitar o acesso.
 - c. Faça download do arquivo dados.zip
 - d. Descompacte a pasta em c:\dados

Virgula-Ponto

OBS: O Genes trabalha somente com ponto.



Virgula-Ponto

OBS: O Genes trabalha somente com ponto.



Ajuda

Finalizar

c:\

- C:\
- dados
- BLUP
- etiquetas
- Google
- Scripts
- ScriptsF
- ScriptsM
- ScriptsR
- Selegen

_FuzzyER. \$\$\$

- aa
- aabbcc.#5c
- aabbcc.3ee
- aabbcc.txt
- aabbcc_t.txt
- acessos.txt
- adair.#4e
- adair.#4y
- adair.4gc
- adair.NOM
- adair1.#e2
- adair1.#e4
- adair1.#e7
- adair1.4gc
- adair1.txt
- Agrupa.#01
- AGRUPA.#03
- AGRUPA.#16
- AGRUPA.#20
- AGRUPA.#6a
- AGRUPA.#8i

✖

Consulte os códigos de erros de execução - on e off line

🩺

Faça diagnóstico de seu arquivo de dados

🕶️

Faça a leitura de seus dados

f

Fale conosco e tire suas dúvidas - acesse o GenesNews no facebook

⚙️

Resolva os problemas de ocx - Instale o setup do bioinfo

√12,25

Veja o símbolo decimal utilizado pelo seu computador

T

Verifique e altere sua trilha de dados

Ok

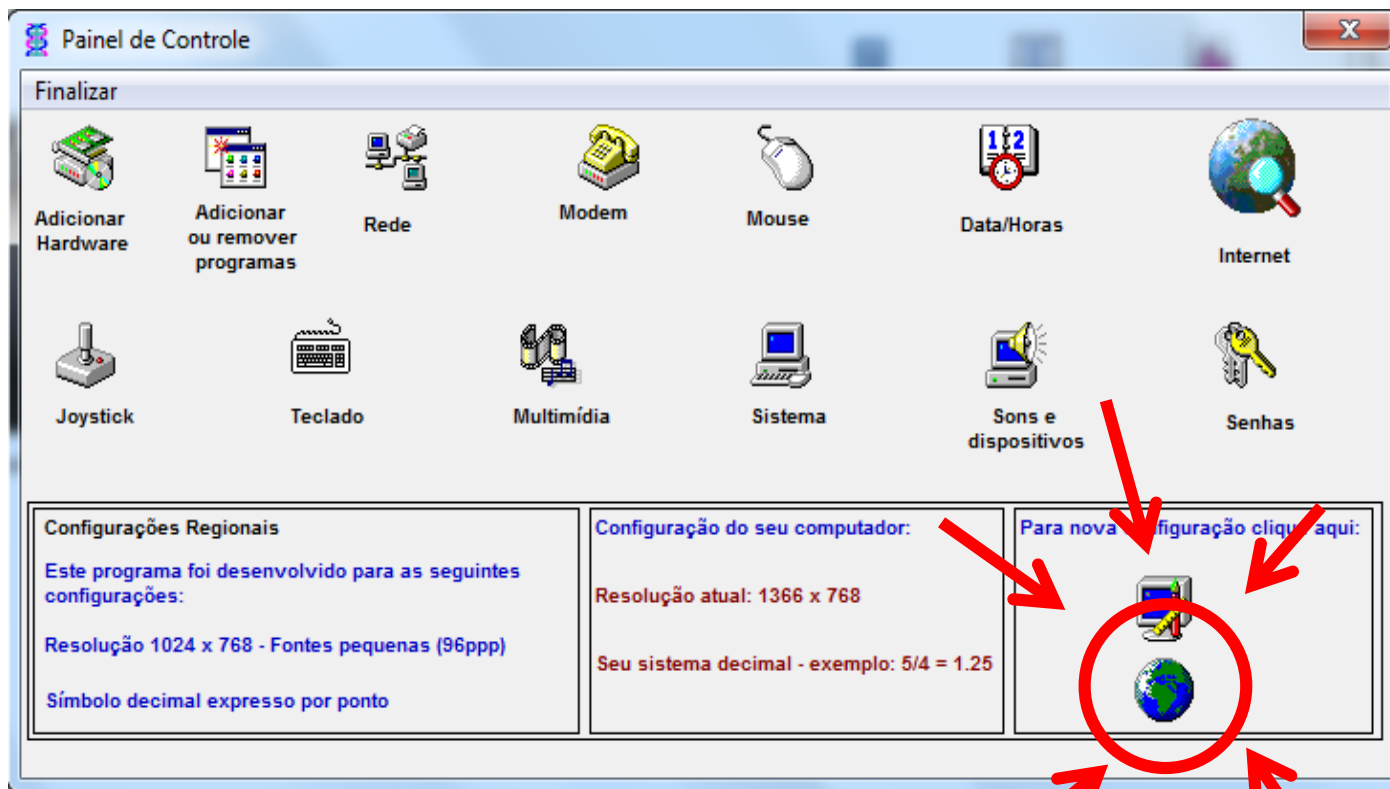
Configurações

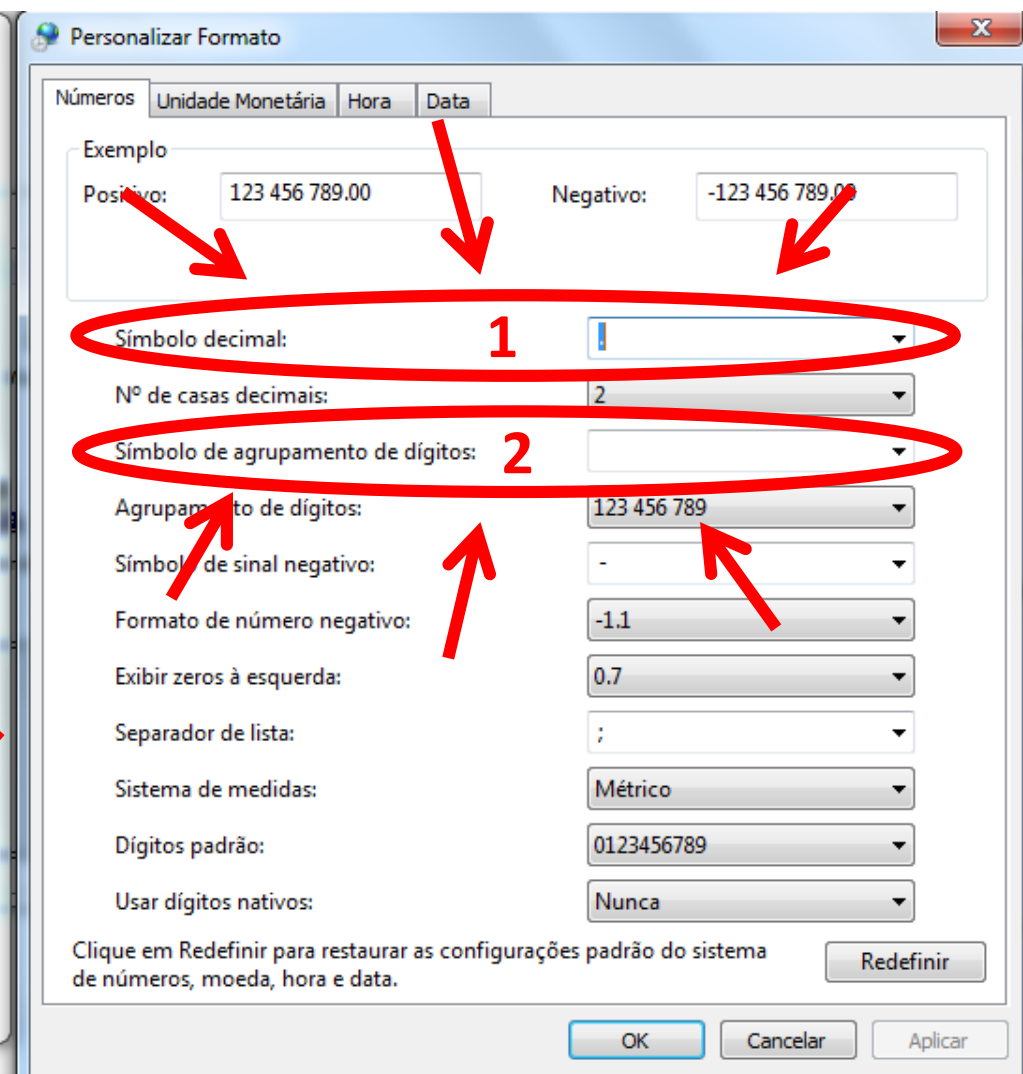
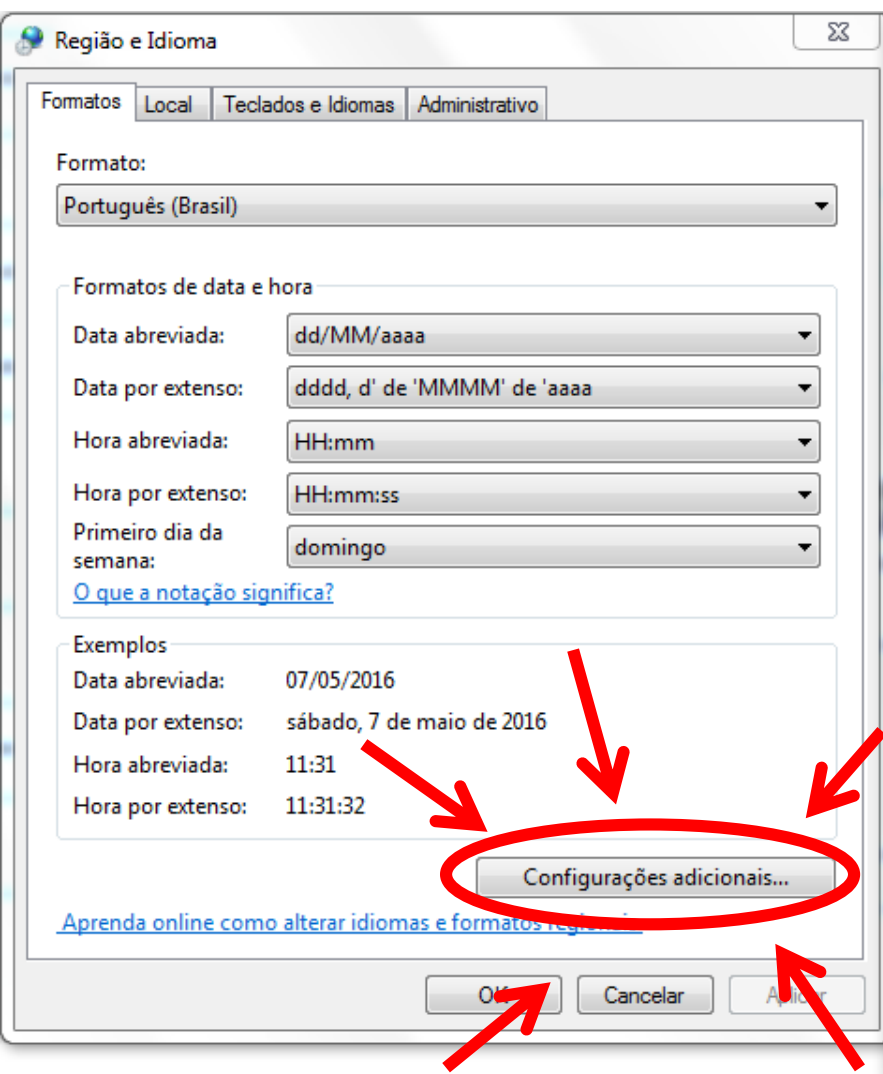


Virgula-Ponto

OBS: O Genes trabalha somente com ponto.

(UTILITÁRIOS -> CONFIGURAÇÕES -> IMAGEM PLANETA -> CONFIGURAÇÕES ADICIONAIS)





1. Símbolo decimal: Colocar ponto (.) como na figura

2. Símbolo de agrupamento de dígitos: Colocar espaço simples () como na figura

Configurações



Trilha de dados (UTILITÁRIOS -> TRILHA DE DADOS)

1. Pasta

- Definição
- Criação

2. Integração



Trilha de Dados

Retornar Padrão

Parâmetros para acesso a arquivo de dados

Usuário Cliente

Trilha C:\dados\

Extensões *.dat;*.txt;*.pm

c:

C:\

dados

BLUP

etiquetas

Google

Scripts

ScriptsF

ScriptsM

ScriptsR

PDF C:\Program Files (x86)\Google\Chrome\Application\chrome.exe ?

R C:\Program Files\R\R-3.4.0\bin\Rscript.exe ?

C:\Program Files\matlab\R2011a\bin\matlab.exe ?

RBIO C:\Users\Cliente\AppData\Roaming\Microsoft\Windows\Start Menu\Programs\Rbio\Rbio. ?

Pessoal C:\Users\Cliente\Dropbox\Eu\

Criar pasta de dados Ok c:\Genes_10-16-2017

Criar atalho no desktop Ok

Roteiro - Configurações

Ao iniciar o programa Genes faça a seguinte checagem:

- a. Veja a definição da trilha de dados (Canto inferior direito da tela). Se não estiver definida, clique na palavra “trilha” e acesse o painel para atualização (para primeira avaliação recomenda-se usar o default que é c:\dados)
- b. Veja a configuração do símbolo decimal. Se estiver com vírgula, passe para ponto. Você pode alterar isto entrando em utilitários e clicando em “configurações”
- c. Faça um teste rodando sua primeira análise. Entre em um procedimento, acesse um conjunto de dados exemplo (clique no ícone representativo de um triângulo amarelo, na tela de entrada do procedimento escolhido)

Primeiras Análises



ANOVA - DIC

(ESTATÍSTICA EXPERIMENTAL -> ANÁLISE DE VARIÂNCIA -> INTEIRAMENTE AO ACASO)

Utilitários	E.Experimental	Biometria	A.Multivariada	Diversidade Genética	Simulação	Matrizes
	Análise de Variância ▶			Alfa látice ▶		
	Anova e Regressão ▶			Inteiramente ao Acaso		
	Estatísticas Descritivas			Tratamentos regulares e não-regulares ▶		
	Imputação de Dados			Blocos Casualizados (DBC) ▶		
	Regressões ▶			Fatorial ▶		
	Correlações ▶			Parcelas Subdividida ▶		
	Comparações entre Médias ▶			Análise de Procedência/Progenie/Planta		
	Métodos de Correção de Estande			Látices ▶		
	Tabela de Contigência			Hierárquico ▶		
	Testes de Qui-quadrado			Redes Neurais Artificiais - DBC		
	Teste de Normalidade					
	Análise Global de Experimentos ▶					
	Análise Não Paramétrica ▶					

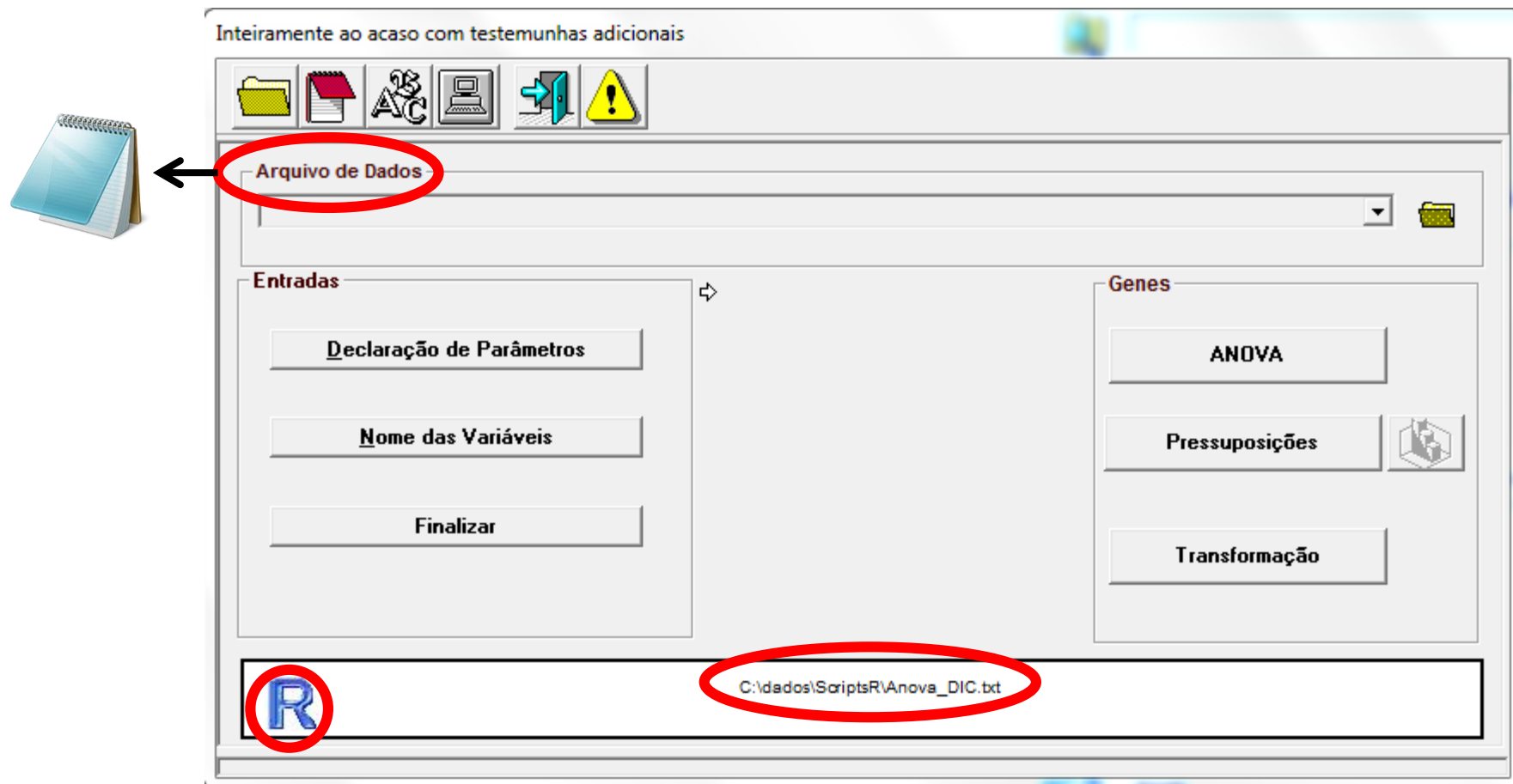


Primeiras Análises



ANOVA - DIC

(ESTATÍSTICA EXPERIMENTAL -> ANÁLISE DE VARIÂNCIA -> INTEIRAMENTE AO ACASO)



Primeiras Análises



Arquivos de Dados

Inteiramente ao acaso com testemunhas adicionais



Arquivo de Dados

Arquivo de Dados

Finalizar

Relatório

Arquivo	c:\dados\Plano1.dat
Número de colunas	7
Número de linhas	30
Linhas em branco	Não há.
Presença de vírgulas	Não há.
Célula em branco	Não há.
Presença de alfanuméricos	Não há.
Números negativos	Não há.
Presença de tabulação	Não há.

Estatísticas

Coluna : 1
Média : 5.5
Número de observações : 30
Variância : 8.534483
Desvio padrão : 2.921384
Máximo : 10
Mínimo : 1



Primeiras Análises



Declaração de Parâmetros

Entradas

Declaração de Parâmetros

Nome das Variáveis

Finalizar

Inteiramente ao Acaso - Testemunhas Adicionais

Arquivo de Dados

c:\dados\Plano1.dat

Parâmetros

Número de Variáveis 5

Número de Genótipos (g) 10

Número de Testemunhas (t) 0

Opções

Retornar

E(QM)

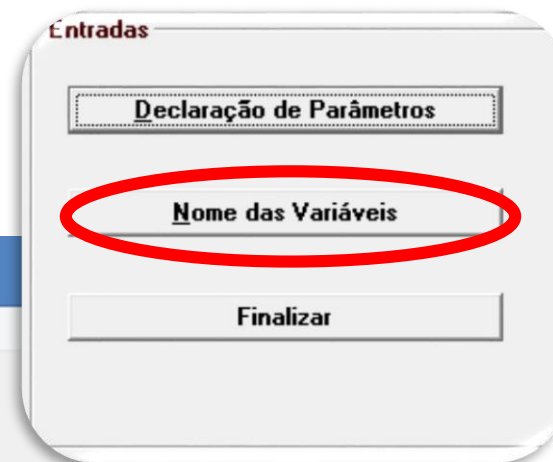
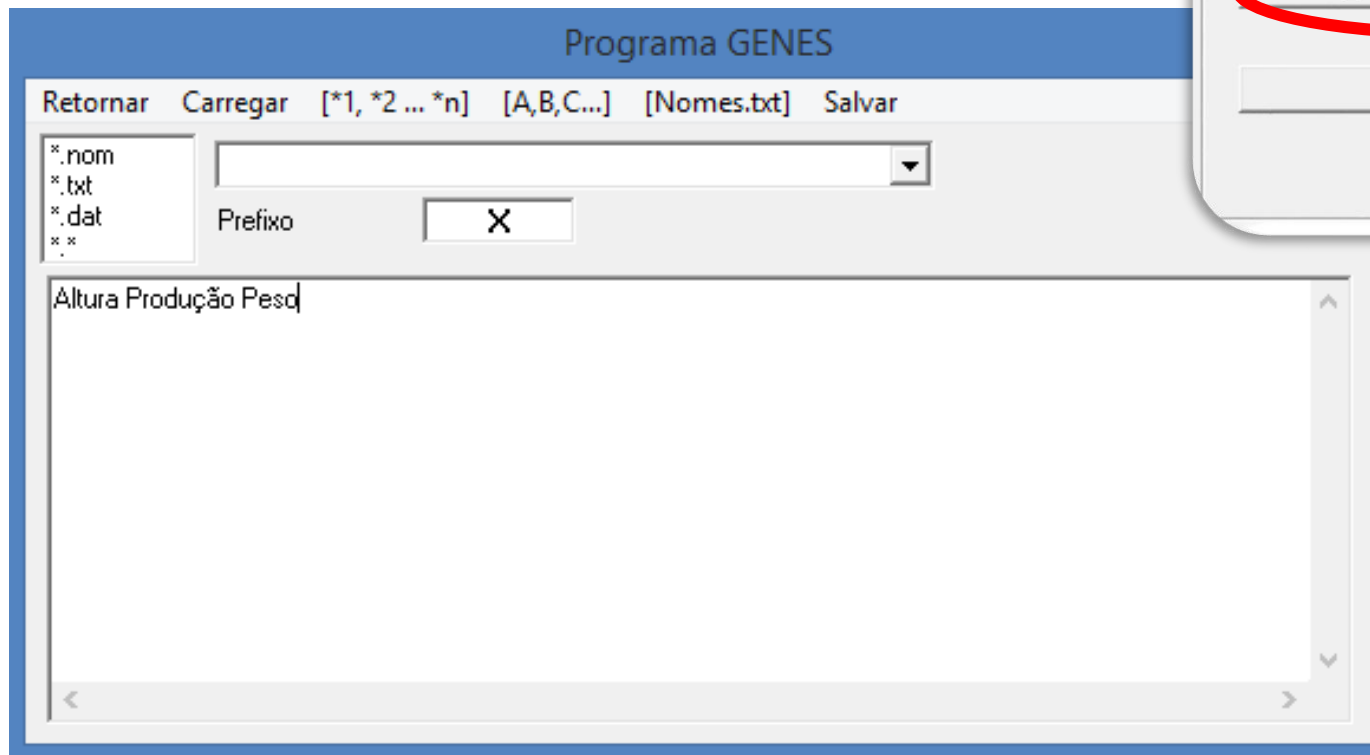
Ler Dados



Primeiras Análises



Nome das Variáveis

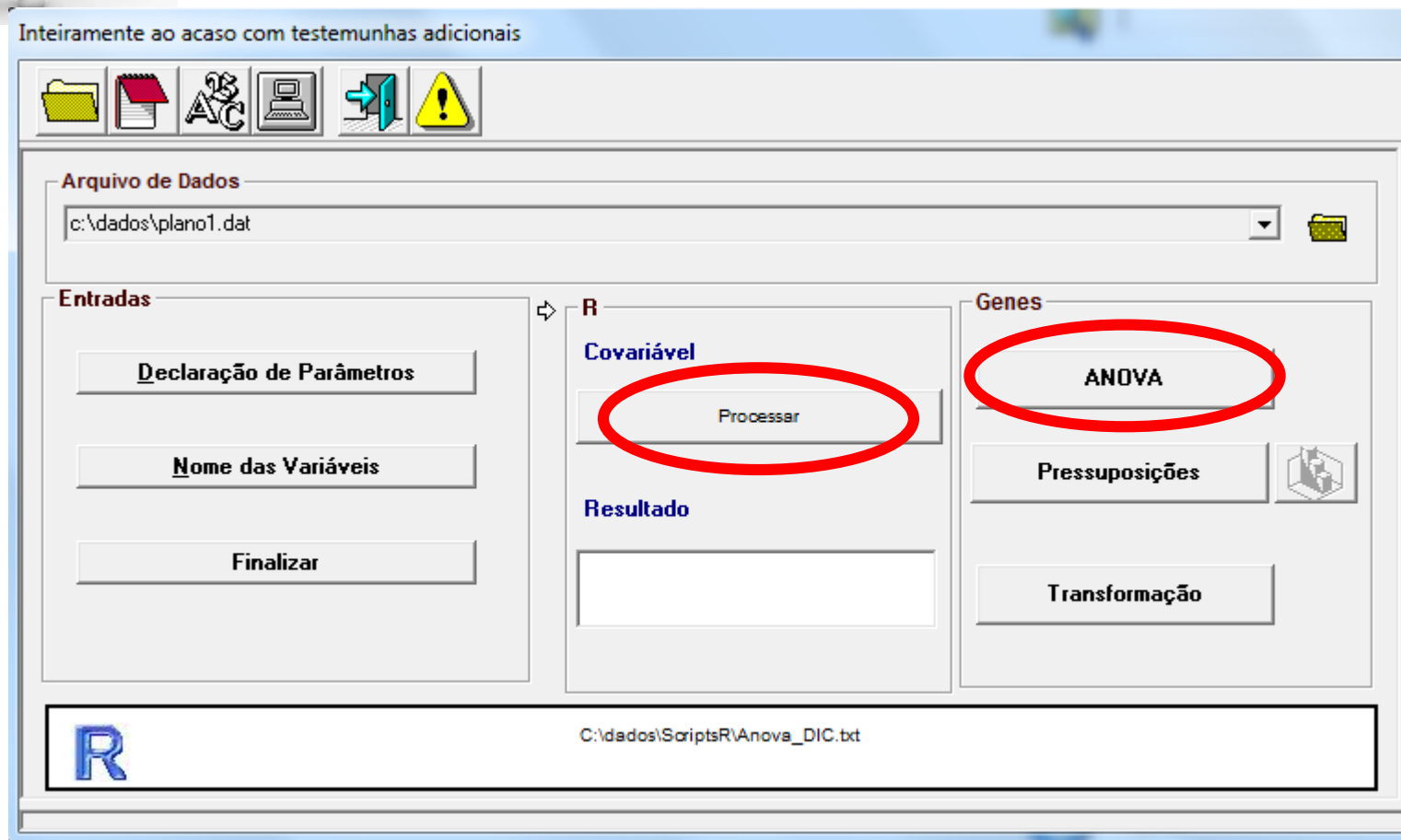


Primeiras Análises



Executar!!

Processar Análise



Erros de Execução

(UTILITÁRIOS -> CÓDIGOS DE ERROS DE EXECUÇÃO)

Erros Rotineiros

1. Utilizar arquivos de dados incompatíveis com as análises
2. Declaração de parâmetros incorretas
3. Presença de tabulação nos dados
4. Linha em branco no final do arquivo
5. Arquivo faltando dados
6. Tabelas com elementos alfanuméricos
7. Análises que requerem o software R e seus pacotes
 - **OBS: A cada vez que instalar uma versão nova do R no computador é interessante executá-lo antes de utilizar esta versão nova nos procedimentos do GENES.**



Primeiras Análises

Processamento dos Dados Via GENES



SAIDA\$\$\$\$.doc

Arquivo Exportar Dados Editar Fonte Cor Frente Cor Fundo Utilitários Calculadoras Finalizar

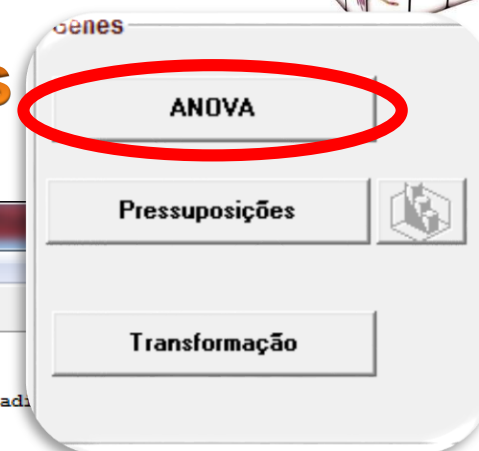
Programa GENES
Arquivo de dados c:\dados\Plano1.dat
Número de variáveis 5
Número de genótipos 10
Número de testemunhas 0
Data 07-26-2016

ANÁLISE DE VARIÂNCIA DA VARIÁVEL => prod

FV	GL	SQ	QM	F	Probabilidade(%)
TRATAMENTOS	9	395801.912137	43977.990237	8.0699	.005628 **
RESÍDUO	20	108992.3526	5449.61763		
TOTAL	29	504794.2647			
MÉDIA geral			524.464333333333		
CV (%)			14.0756045411393		

PARÂMETROS GENÉTICOS E AMBIENTAIS- Obtidos para genótipos

VARIÂNCIA FENOTÍPICA (média)	14659.330079
VARIÂNCIA AMBIENTAL (média)	1816.53921
VARIÂNCIA GENOTÍPICA (média)	12842.790869
HERDABILIDADE (US: média da família) - %	87.6083
CORRELAÇÃO INTRACLASSE (US: parcela)- %	70.2083
COEFICIENTE DE VARIAÇÃO GENÉTICO (%)	21.608
RAZÃO CVg/CVe - Média dos genótipos	1.5351
NÚMERO MÉDIO DE REPETIÇÕES	3.



Primeiras Análises



Processamento Matrizes Disponíveis



Matrizes Disponíveis

Processa Retorna Marcar todas

Variâncias e Covariâncias

☐ Fenotípicas (Média) C:\Users\Vinicius\Desktop\Curso GENES\Aula_01\Exemplos\dadoscfe.d

☐ Genotípicas C:\Users\Vinicius\Desktop\Curso GENES\Aula_01\Exemplos\dadosoge.c

☐ Residuais C:\Users\Vinicius\Desktop\Curso GENES\Aula_01\Exemplos\dadosre.d

Correlações

☐ Fenotípicas C:\Users\Vinicius\Desktop\Curso GENES\Aula_01\Exemplos\dadosrfe.ds

☐ Genotípicas C:\Users\Vinicius\Desktop\Curso GENES\Aula_01\Exemplos\dadosrge.d

☐ Residuais C:\Users\Vinicius\Desktop\Curso GENES\Aula_01\Exemplos\dadosrre.ds

Médias e Variâncias

☐ Médias dos Genótipos (inclui o número de repetições)
C:\Users\Vinicius\Desktop\Curso GENES\Aula_01\Exemplos\dadosmex.dat

☐ Médias dos Genótipos (não inclui o número de repetições)
C:\Users\Vinicius\Desktop\Curso GENES\Aula_01\Exemplos\dadosmed.dat

☐ Variância dentro de genótipos (inclui o número de repetições)
C:\Users\Vinicius\Desktop\Curso GENES\Aula_01\Exemplos\dadosvar.dat

Primeiras Análises



Processamento dos Dados Via Integração R

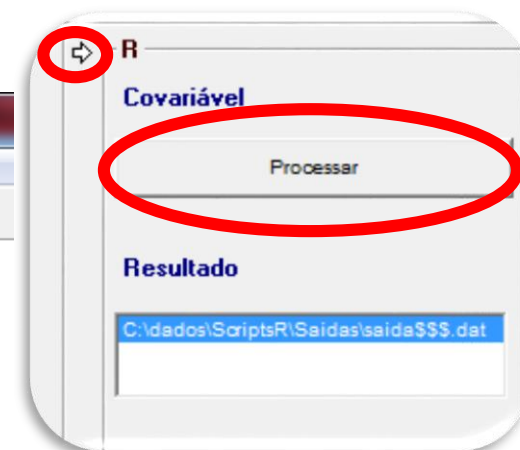
```

C:\dados\ScriptsR\Saidas\saida$$$.dat
Arquivo  Exportar Dados  Editar  Fonte  Cor Frente  Cor Fundo  Utilitários  Calculadoras  Finalizar
[Icons]
*****
**                                     **
**      Programa Genes - Programa R      **
**                                     **
**  Procedimento: Anova DIC              **
**                                     **
*****

Informações Preliminares
      Trat      Rep      prod      acam      alt
Min.   : 1.0   Min.   :1   Min.   :227.6   Min.   : 59.78   Min.   :270.9
1st Qu.: 3.0   1st Qu.:1   1st Qu.:461.0   1st Qu.:357.25   1st Qu.:305.3
Median : 5.5   Median :2   Median :561.3   Median :452.74   Median :394.7
Mean   : 5.5   Mean   :2   Mean   :524.5   Mean   :441.59   Mean   :382.7
3rd Qu.: 8.0   3rd Qu.:3   3rd Qu.:618.4   3rd Qu.:529.27   3rd Qu.:415.7
Max.   :10.0   Max.   :3   Max.   :756.1   Max.   :862.62   Max.   :566.6

      flor      nesp
Min.   :226.5   Min.   : 14.92
1st Qu.:274.8   1st Qu.:257.00
Median :319.7   Median :394.04
Mean   :328.0   Mean   :380.85
3rd Qu.:363.2   3rd Qu.:496.62
Max.   :556.3   Max.   :851.41

Anova - Variável = prod
      Df Sum Sq Mean Sq F value    Pr(>F)
Trat    9 395802   43978    8.07 5.62e-05 ***
Residuals 20 108992    5450
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Teste Normalidade = prod
  
```





GENES - Básico (UTILITÁRIOS)

1. Calculadora on line
2. Estatística on line
3. GenesCalc
4. Matriz
5. Raízes de polinômio
6. Planilha de Experimentos
7. Tabelas de Probabilidade
8. Funções de Probabilidade

Atualização- Versão atual : 1990.2017.18

Pessoal

Cruz, C.D. GENES - a software package for analysis in experimental statistics and quantitative genetics. Acta Scientiarum. v.35, n.3, p.271-276, 2013

Cruz, C.D. Genes Software – extended and integrated with the R, Matlab and Selegen. Acta Scientiarum. v.38, n.4, p.547-552, 2016

Rede social

Tutoriais

Perguntas Mais Frequentes

Soluções e Configurações

Análise e processamento de dados pelo software R (Link)

FIM – AULA 01

DUVIDAS?

