Manipulação de dados genômicos: de 3K a dados de sequência

Ricardo Ventura
UoGuelph – Canada
Beef Improvement Opportunities – Canada
FZEA USP - Pirassununga

Flávio Schenkel – UoGuelph - Canada
Steve Miller – AgResearch – New Zealand

BIO

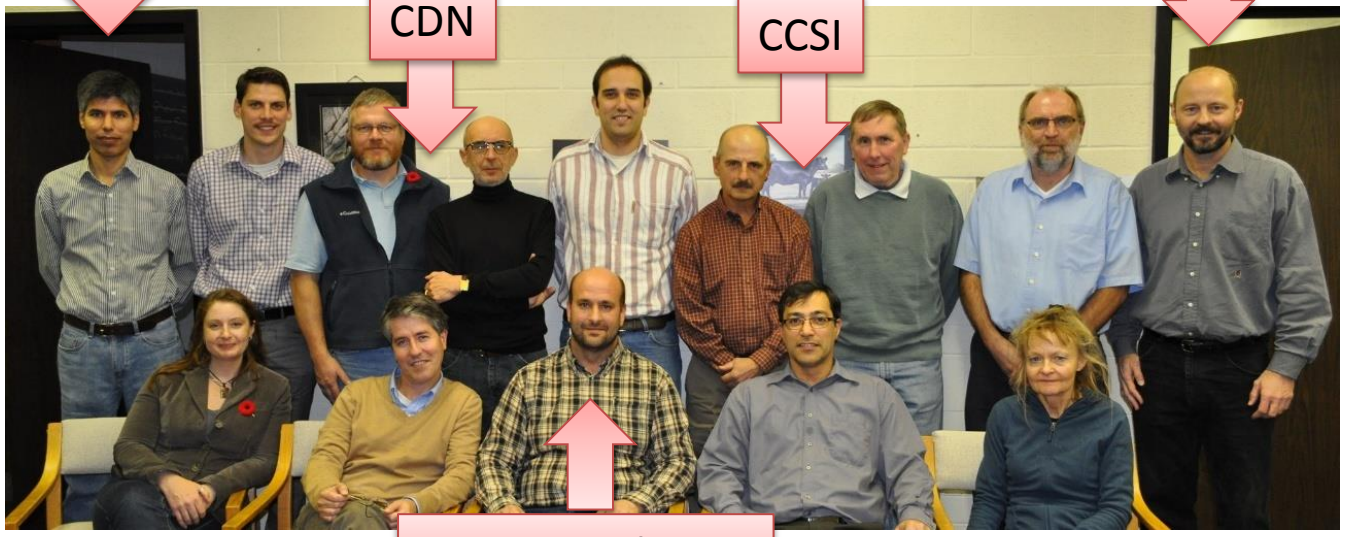
SEMEX

CGIL Director

New model: UoG + Industry

CDN

CCSI



AgrResearch - NZ

bio
Bridging Intelligence

BIO:
Genomicist

Tópicos:

Manipulação de dados genômicos
- de 3K a dados de sequenciamento

Que linguagem utilizar: C++, Fortran, Perl, SAS, R? Excel?

Banco de dados genômicos: Lições e novos projetos

Compartilhamento de genótipos entre grupos/países

Conclusões

Qual linguagem utilizarei?

C++, Fortran, Perl, SAS, R? Excel?

```
mov ax,4       // (I/O Func.)
mov bx,1       // (Output func)
lds cx,“EMBRAPA, obrigado pelo convite!" // (address of the string)
mov dx,6       //  (length of the string)
int 0x21       // system call
```

```
print ( "EMBRAPA, obrigado pelo convite!")
```

```
#!/usr/bin/perl
print "EMBRAPA, obrigado pelo convite!";
```

```
public class Hello {
    public static void main(String[] args) {
        System.out.println(“EMBRAPA, obrigado pelo convite!");
    }
}
```

```
#include <iostream>
using namespace std;

int main()
{ cout << "Obrigado pelo convite EMBRAPA!" << endl;
 system("pause"); }
```

```
print ( "Obrigado pelo convite EMBRAPA!")
```

```
01001111 01100010 01110010 01101001 01100111 01100001
01100100 01101111 00100000 01110000 01100101 01101100
01101111 00100000 01100011 01101111 01101110 01110110
01101001 01110100 01100101 00100000 01000101 01001101
01000010 01010010 01000001 01010000 01000001 00100001
```

Linguagem compilada:

• Converte uma linguagem de alto nível do código fonte legível de um programa em um **programa executável** em linguagem de máquina

• Converte todo um código de fonte de uma só vez

• Compilação deve ser finalizada **antes de executar** o programa

•Todo o processo de compilação é **simplesmente** usado para **criar** algo que possa ser **executado**.

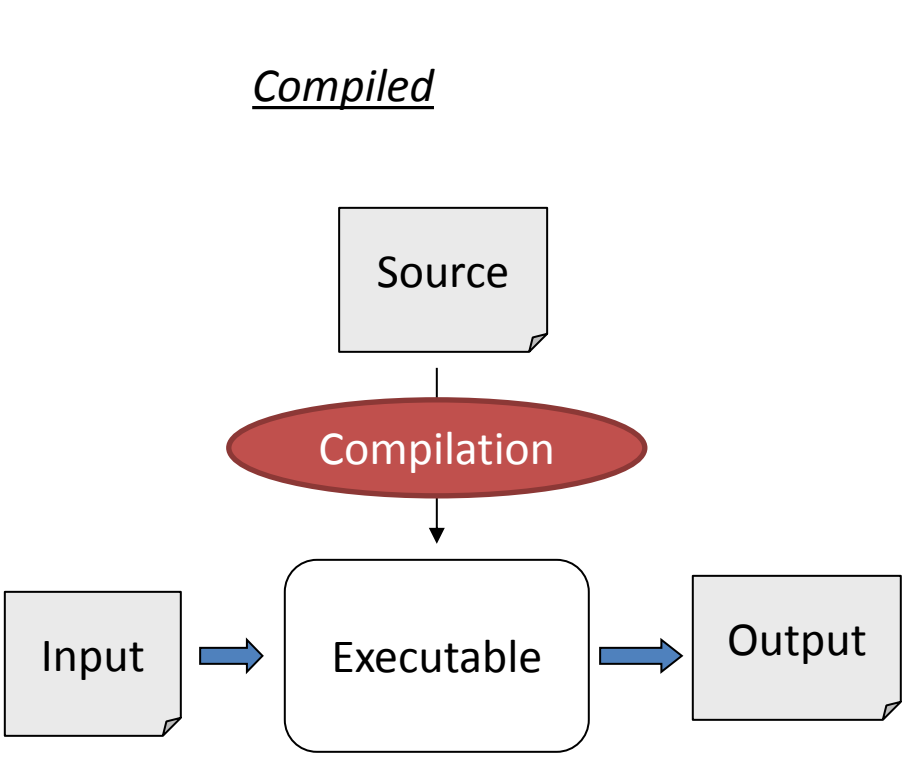
Linguagem Interpretada:

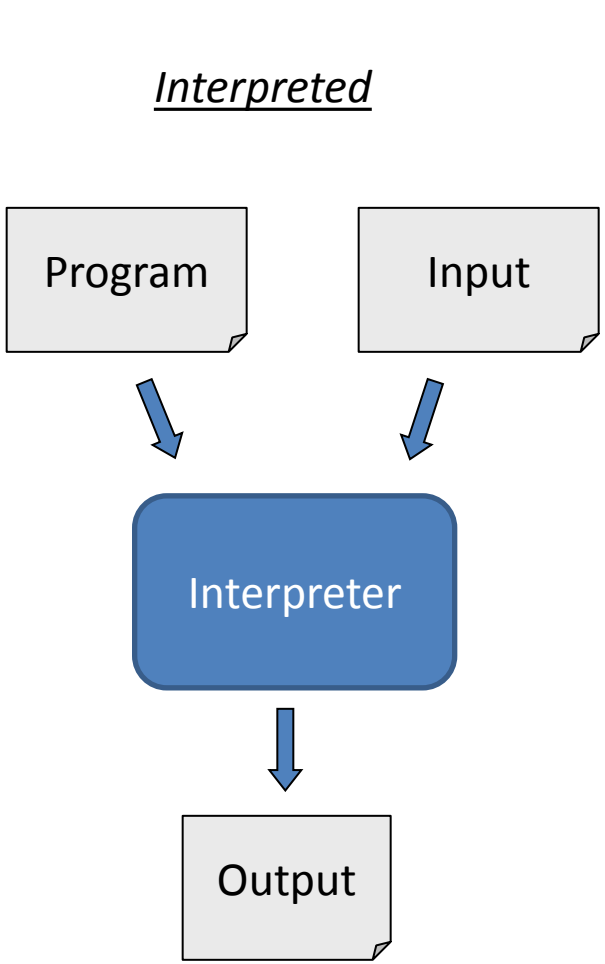
código fonte é executado **diretamente** por um interpretador

Variáveis criadas no programa será mantido em uma espécie de tabela de pesquisa que mapeia os nomes de seus dados.

Se você **executar o programa novamente**, o intérprete terá que **gerar o código novamente.**

Linguagens Interpretadas e compiladas





Exemplos de linguagem compilada:
Fortran 90, C, C + +, Pascal

Exemplos de linguagem interpretada:
Java Script, Perl, Phyton, R

Normalmente mais lentos e mais flexíveis

Exemplo de linguagem híbrida:
(*) Java:
compilada e **interpretada por uma máquina JVM**

Quais critérios para escolha?

Qual o tipo de terreno?

Que velocidade quero atingir?

Qual a distância irei percorrer?

Qual a frequência de utilização (cada caso)?

Quais critérios para escolha da linguagem?

Qual a frequência de utilização?

Quais Sistema Operacional? VM?

Qual configuração do servidor?

Qual o tempo de aprendizado?

Gráficos? Dados em larga escala?

Receita mineira para esse problema!

“Mexidão de linguagens” complementares

Ricardo e uso de linguagens de programação

Antes do envolvimento com genômica

2006	2007	2008	2009	2010
Java SAS	Java Html / Java Script SAS	JSTF SAS	Matlab SAS	SAS Excel

Após genômica

2010	2011	2012	2013	2014
Excel C++	Excel C++ R	Excel C++	Excel C++ Perl	Awk / Sed R C++ Perl

C++ ou Linux (AWK, Sed …)

CONTAR NUMERO DE LINHAS NO ARQUIVO

wc -l file.txt

```
Ifstream file;
file.open(“genotype.txt”);
string stringLine = “”;
int countLines = “”;
if (file.is_open())
 {
        while(file >> markerName)
        {
              getline(file, stringLine);
               countLines ++;
        };
 }
      else
        cout << "Problems to open file genotype.txt " << endl;
cout << “Number of lines: “ << countLines << endl;
file.close();
```

Comandos Linux, AWK e Sed …

CALCULAR NUMERO DE COLUNAS LINUX

cat PR_26_OUT_INVERTIDO_CHECAGEMRESULTBEAGLE.txt | awk '{ print NF}'

TROCAR TAB POR SPACE:

awk '{ gsub(/\t/, " "); print }' < AB_5032_1TO5032.txt > output.txt

FOLDER ONDE O PROCESSO ESTA RODANDO:

ls -l /proc/1175/cwd ..... "1175" e' o numero do processo PID

CONTAR NUMERO DE LINHAS NO ARQUIVO

wc -l file.txt

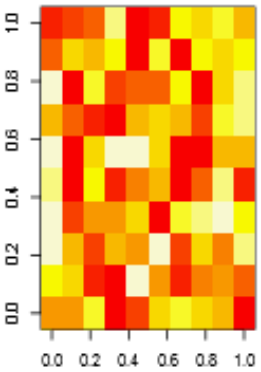
OBTER NUMERO DE MARCADORES DO ARQUIVO GEBV - COLUNA DE GENOTIPOS

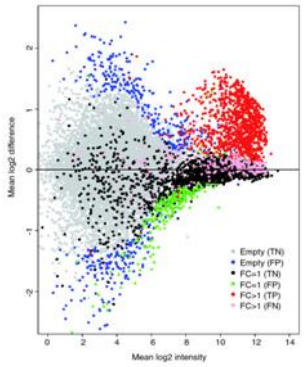
awk ' {print length($2)}' genoHD_05.txt

INSERIR ESPACO ENTRE GEBV GENOTYPE FORMAT FOR R!

awk 'NR >1 { gsub(/./, "& ", $2); print $1 " " $2; }' TESTE > teste.txt

- Ampla variedade de pacotes estatísticos
- **Técnicas gráficas**
- Altamente extensível
- C, C + + e Fortran acionados em tempo de execução
- Utilizar R dentro de rotinas C e C++





Linguagem: cálculos matemáticos (compiladas)

Tarefas com alta repetitividade
Desenvolvimento de métodos

Paradigma procedural

Direcionado a funcionalidades

Linguagem de baixo nível

C Language

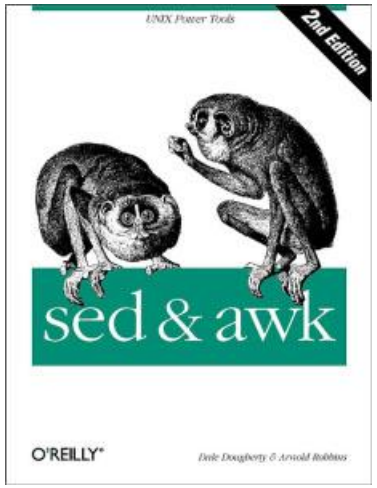
Multi-paradigma (Procedural + Orientação objetos

Direcionado a objetos

Linguagem de nível intermediário

Linguagens de script (interpretadas!)

Tarefas de baixa repetitividade
Entrada e saída de dados no BD
Manipulação de sequências



Como ler/manipular arquivos genótipos?

R?

AWK ?

C++ ?

Inserir em BD e buscar subset?

Carregar todos os dados na memória?
Até 50K esse erro ainda pode ser cometido!

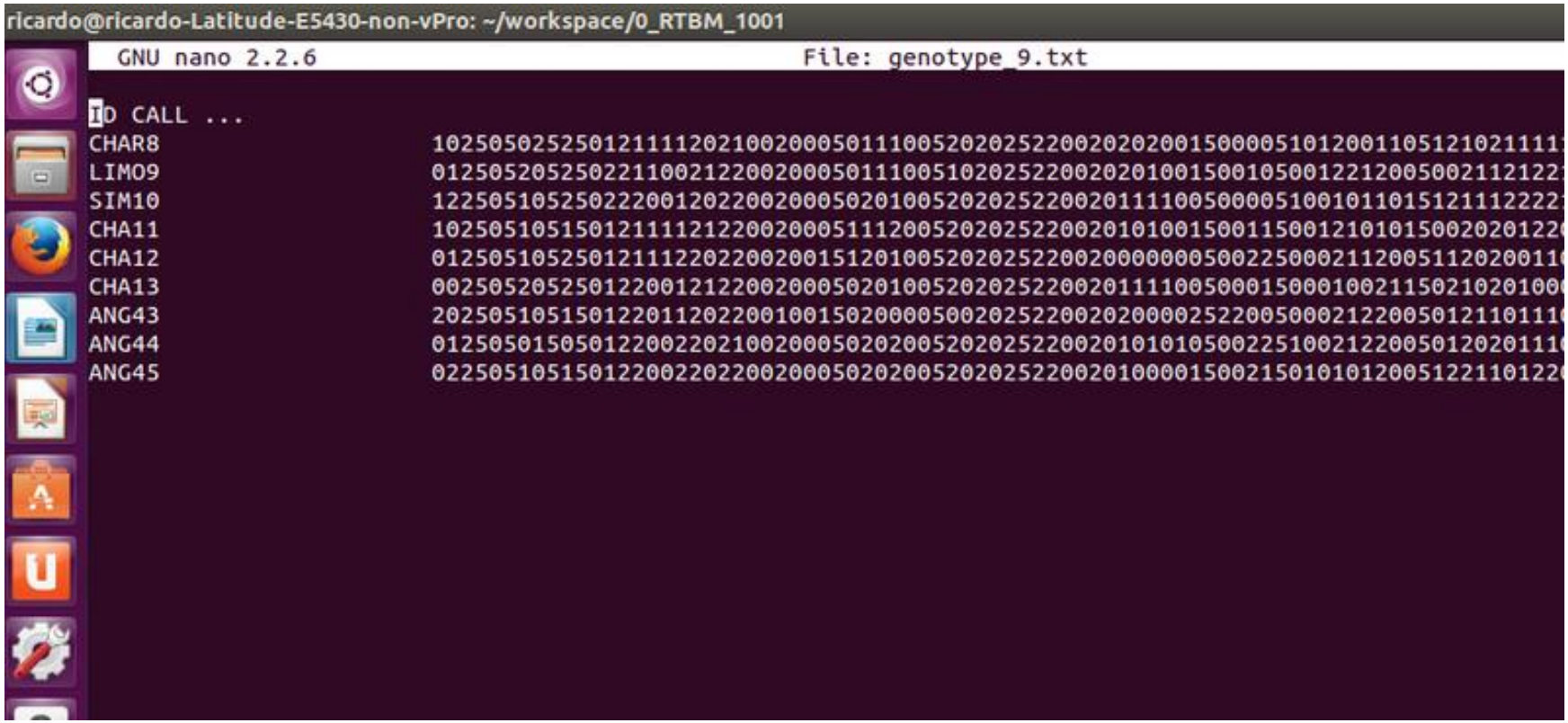
```
[Header]
GSGT Version      1.9.4
Processing Date   8/31/2012 5:20 PM
Content           BovineSNP50_v2_C.bpm
Num SNPs     54609
Total SNPs   54609
Num Samples 75
Total Samples     75
[Data]
SNP Name     Sample ID   Allele1 - Forward Allele2 - Forward Allele1 - Top     Allele2 - Top     Allele1 - AB      Allele2 - AB
      Y
ARS-BFGL-BAC-10172       X439Y G     G     G     G     B     B     0.9506       0.027 1.000
ARS-BFGL-BAC-1020 X439Y -     -     -     -     -     -     0.0000       0.177 0.630
ARS-BFGL-BAC-10245       X439Y C     C     G     G     B     B     0.7579       0.160 1.424
ARS-BFGL-BAC-10345       X439Y A     C     A     C     A     B     0.9276       0.679 0.870
ARS-BFGL-BAC-10365       X439Y G     G     C     C     B     B     0.5335       0.076 1.188
ARS-BFGL-BAC-10375       X439Y A     G     A     G     A     B     0.9567       0.372 0.418
ARS-BFGL-BAC-10591       X439Y A     G     A     G     A     B     0.9003       0.311 0.481
ARS-BFGL-BAC-10867       X439Y G     C     C     G     A     B     0.9434       0.428 0.468
ARS-BFGL-BAC-10919       X439Y A     G     A     G     A     B     0.8526       0.876 0.982
ARS-BFGL-BAC-10951       X439Y T     T     A     A     A     A     0.5140       0.601 0.034
ARS-BFGL-BAC-10952       X439Y A     A     A     A     A     A     0.9512       1.331 0.033
ARS-BFGL-BAC-10960       X439Y -     -     -     -     -     -     0.0000       0.047 0.241
ARS-BFGL-BAC-10972       X439Y G     C     C     G     A     B     0.8759       1.016 0.877
ARS-BFGL-BAC-10975       X439Y A     G     A     G     A     B     0.7379       0.525 0.443
ARS-BFGL-BAC-10986       X439Y G     G     C     C     B     B     0.9309       0.045 0.877
ARS-BFGL-BAC-10993       X439Y T     C     A     G     A     B     0.9014       1.087 0.986
ARS-BFGL-BAC-11000       X439Y T     T     A     A     A     A     0.9686       0.578 0.023
ARS-BFGL-BAC-11003       X439Y T     T     A     A     A     A     0.9215       1.576 0.060
ARS-BFGL-BAC-11007       X439Y T     C     A     G     A     B     0.9454       0.689 0.695
ARS-BFGL-BAC-11025       X439Y -     -     -     -     -     -     0.0000       0.171 0.808
ARS-BFGL-BAC-11028       X439Y A     A     A     A     A     A     0.9678       0.427 0.019
ARS-BFGL-BAC-11034       X439Y T     C     A     G     A     B     0.9509       0.591 0.696
ARS-BFGL-BAC-11039       X439Y C     C     G     G     B     B     0.9658       0.017 0.982
ARS-BFGL-BAC-11042       X439Y G     G     G     G     B     B     0.8506       0.019 1.305
ARS-BFGL-BAC-11044       X439Y C     C     G     G     B     B     0.9654       0.022 1.332
```

Diferente numero de colunas comparado ao apresentado pelo Prof. Fabyano

Na verdade existem dezenas versoes diferentes

```
ARS-BFGL-BAC-11152       X439Y G     G     G     G     B     B     0.2517       0.114 0.829
ARS-BFGL-BAC-11154       X439Y T     C     A     G     A     B     0.8469       1.099 1.270
ARS-BFGL-BAC-11193       X439Y A     A     A     A     A     A     0.9748       1.201 0.008
ARS-BFGL-BAC-11200       X439Y T     T     A     A     A     A     0.6973       0.373 0.035
ARS-BFGL-BAC-11206       X439Y G     G     G     G     B     B     0.9139       0.092 1.037
ARS-BFGL-BAC-11215       X439Y A     A     A     A     A     A     0.4502       0.327 0.058
ARS-BFGL-BAC-11218       X439Y A     A     A     A     A     A     0.9363       1.185 0.026
ARS-BFGL-BAC-11276       X439Y A     A     A     A     A     A     0.9426       1.174 0.027
ARS-BFGL-BAC-11281       X439Y G     G     G     G     B     B     0.8449       0.120 1.251
ARS-BFGL-BAC-11283       X439Y A     A     A     A     A     A     0.8673       2.176 0.011
ARS-BFGL-BAC-11284       X439Y A     A     A     A     A     A     0.9036       1.206 0.150
ARS-BFGL-BAC-11290       X439Y G     G     C     C     B     B     0.8221       0.093 1.673
```

Recodificando arquivo de genótipos



AA = 2 / BB = 0 / AB = 1 / BA = 1 / Missing = 5

Leitura de arquivos contendo dados genômicos

Ex: 35.000 animais genotipados

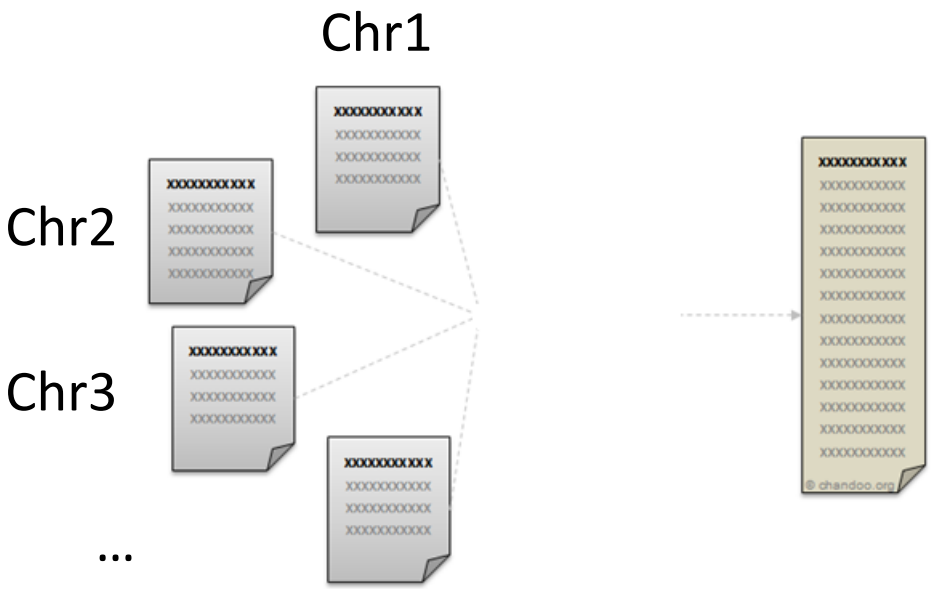
Uso de subset contendo 5000 animais

```
Ifstream file;
file.open("genotype.txt");
string genotype = "";
if (file.is_open())
 {       while(file >> markerName)
             {
                 getline(file, genotype);
                 currentAnimal.geno = genotype ;
             };
 }
      else
             cout << "Problems to open file genotype.txt " << endl;
file.close();
```

Filtros ( ) ...

Execução () ...

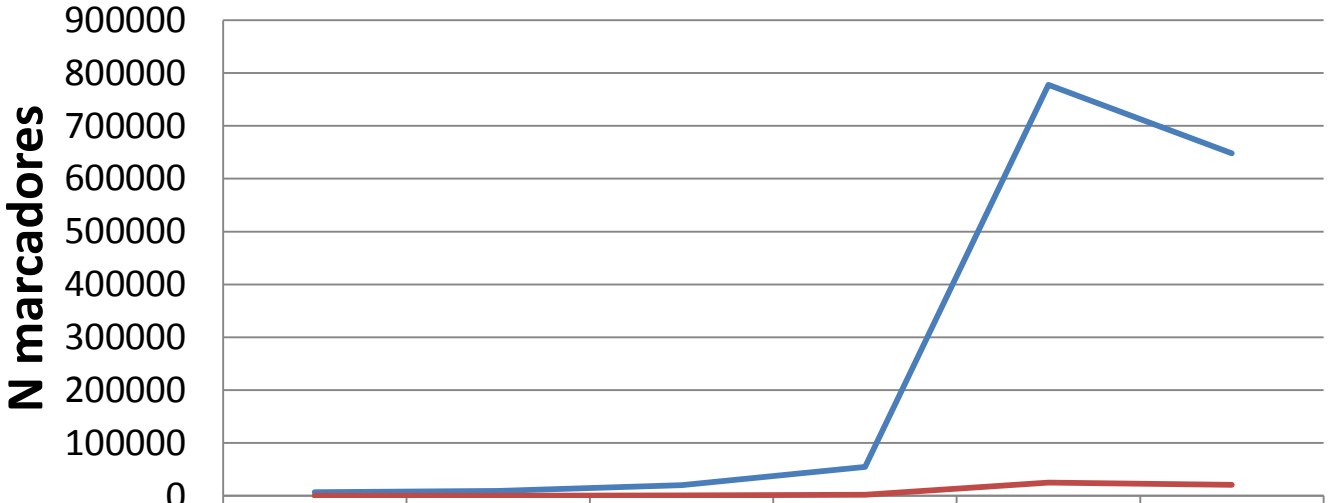
Como ler/manipular arquivos genótipos?



OMP!

Criação de **arquivos temporários** binários

Painél



	6K	9K	20K	50K	777K	648K
Total	6909	9000	20000	54609	777962	648000
Chr	222,87	290,32	645,16	1761,6	25096	20903

Sequência: ~ 13.000.000 = 420.000 / chr = ~ 1 painél HD / chr

OMP – C, C++ and Fortran

```c
/* C/C++ Example */

#pragma omp parallel for
private(temp) shared(n,a,b,c)
{
  for(i=1; i<=n; i++){
    temp = 2.0*a[i];
    a[i] = temp;
    b[i] = c[i]/temp;
  }
}
```

```
!Fortran example

!$omp parallel do private(temp)
shared(n,a,b,c)
do i = 1, n
  temp = 2.0*a(i)
  a(i) = temp
  b(i) = c(i)/temp
Enddo
```

Montahaplotipo = Montahaplotipo + string;

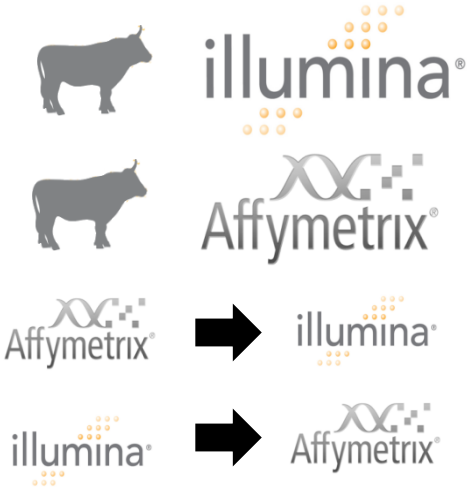
OMP – C, C++ and Fortran

Divisão de tarefas por cromossomo ou indivíduo:

Exemplos:

- Conversão de dados
- Cálculo da matriz EDM
- Descoberta de paternidade

Combinando painéis HD Illumina e Affymetrix



0100100111000100111222002020212101211100200

2251102211221000122221021021 2222

5555105555555550155555555555255555555555555

555510555555555015555555555552555

C++: Necessidade de “debug” e códigos detalhados!
Sincronizar 2 mapas (Illumina e Affymetrix)
Divisão por Chr para evitar problemas de memória
substring();

Painéis comerciais: Affymetrix e Illumina

50K V1 – 54.001 SNPs

50K V2 – 54.609 SNPs

Affymetrix®

10K

648K

3K – 2.900 SNPs

6K – 6.909 SNPs

6K GGP

Angus GGP-HD (79.000)/ GGP – LD (10.000) / GGP – HD (80.000)

Como filtrar SNPs/Animais de diversos arquivos?

Como recodificar os arquivos vindos do laboratório?

Manter arquivos originais após processo de recodificação?

Como armazenar animais genotipados oriundos de diferentes painéis?

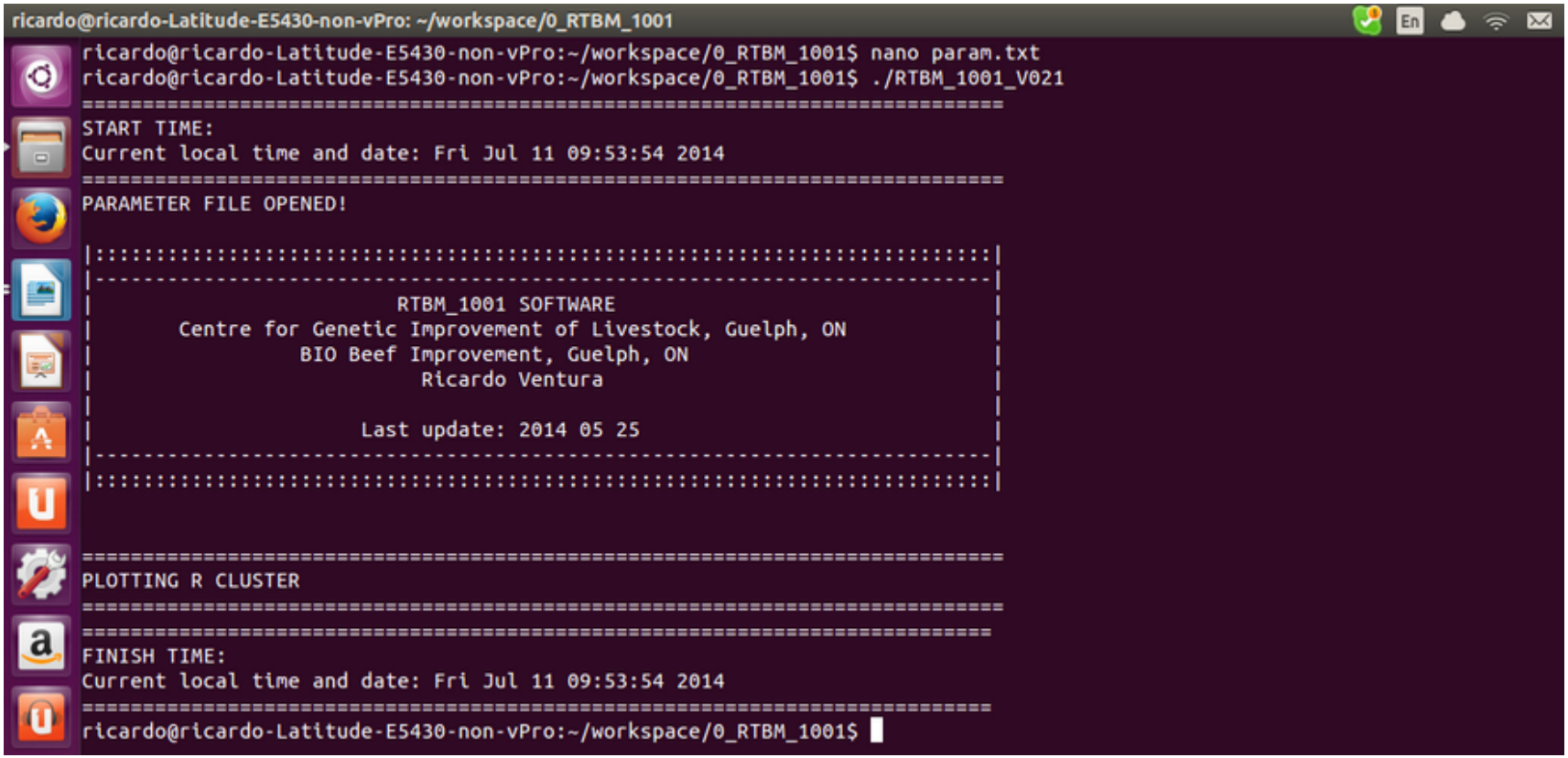
O que ocorre se mudar a posição dos SNPs no mapa?

Como identificar diferentes tipos de amostras/painéis em um único ID?

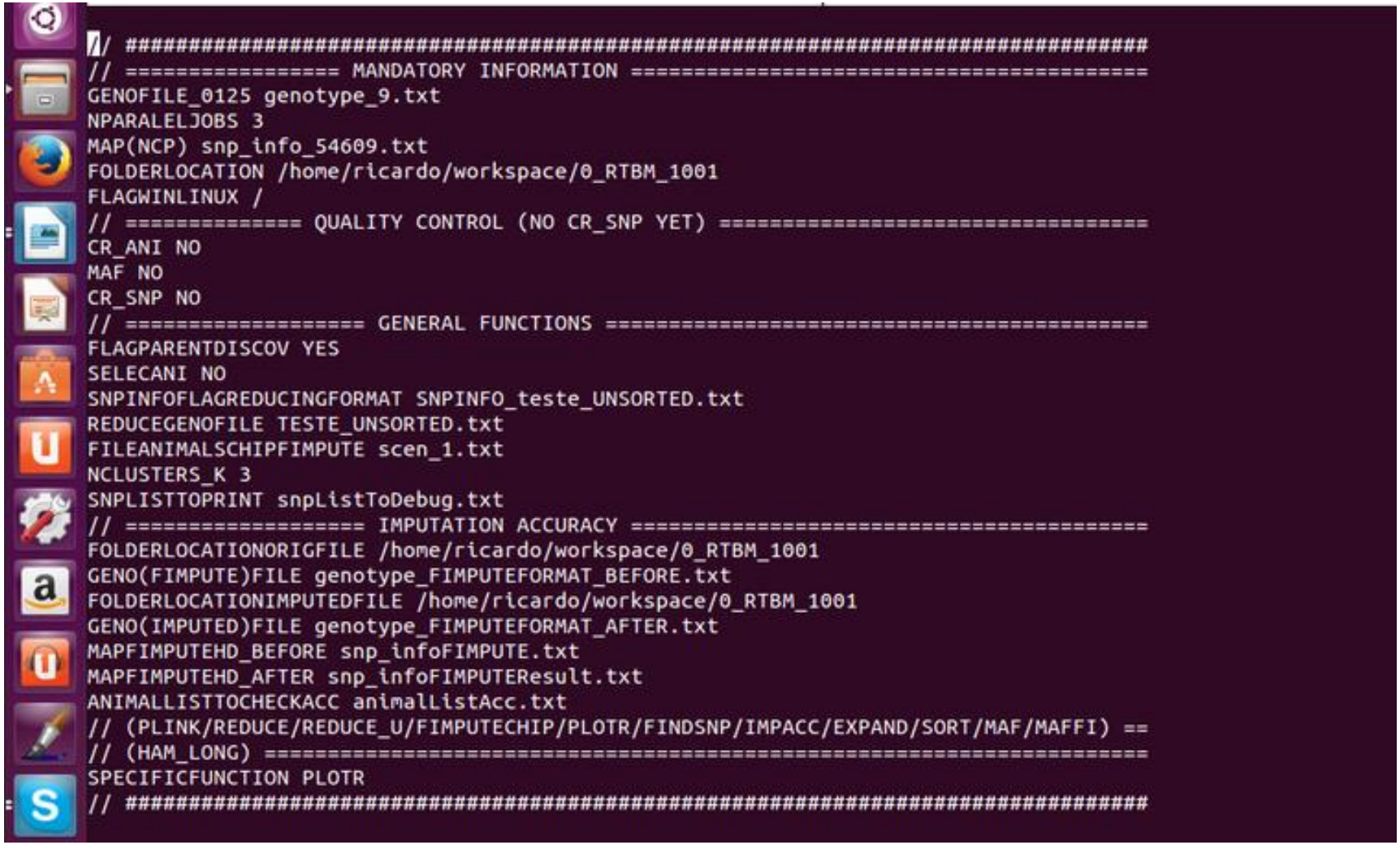
Como armazenar genótipos de animais imputados?

Backup?

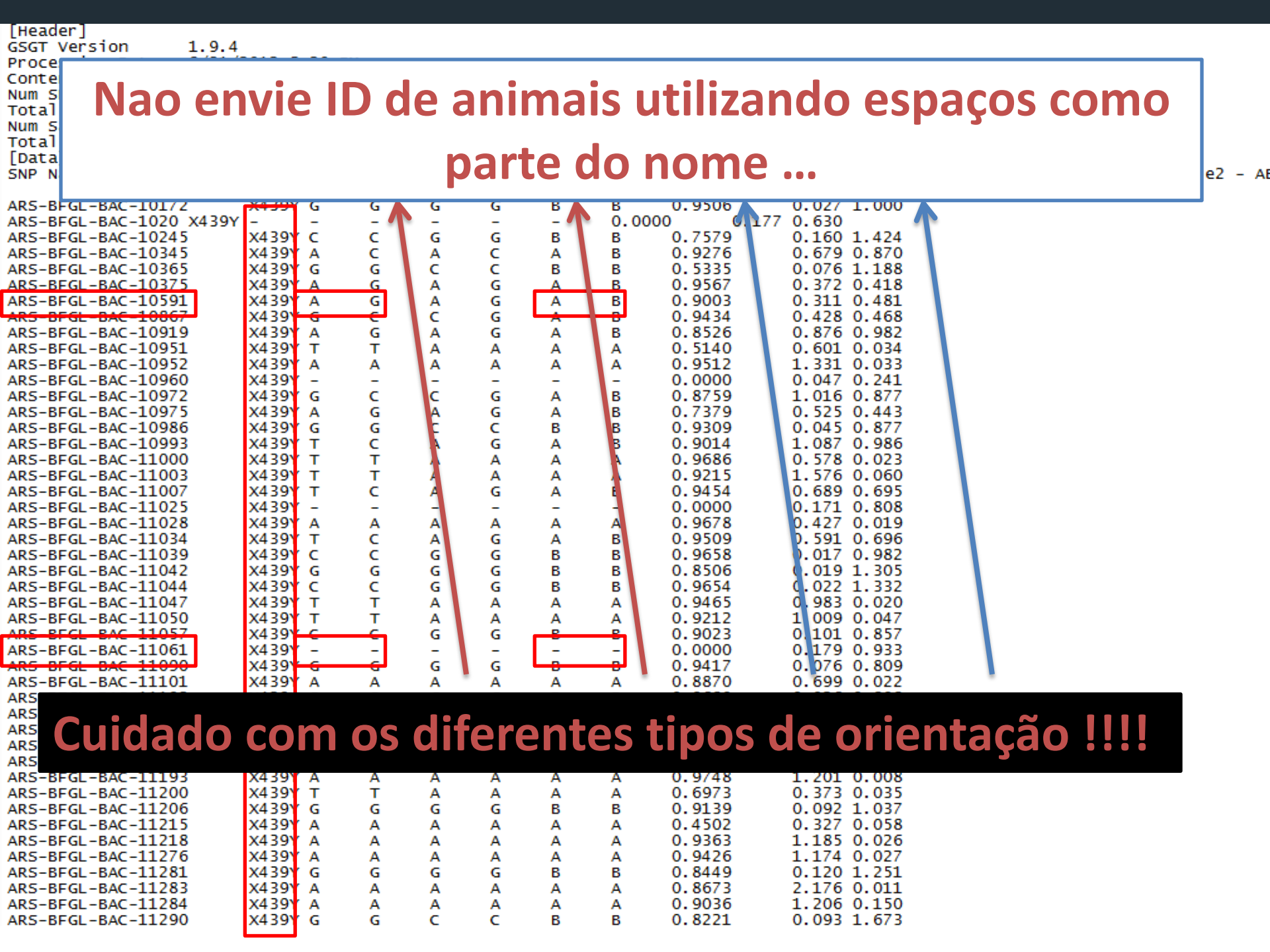
Centralizando rotinas em um único software



Centralizando rotinas em um único software



```
// ##############################################################################
// ================ MANDATORY INFORMATION ======================================
GENOFILE_0125 genotype_9.txt
NPARALELJOBS 3
MAP(NCP) snp_info_54609.txt
FOLDERLOCATION /home/ricardo/workspace/0_RTBM_1001
FLAGWINLINUX /
// ============== QUALITY CONTROL (NO CR_SNP YET) ==============================
CR_ANI NO
MAF NO
CR_SNP NO
// ================== GENERAL FUNCTIONS ========================================
FLAGPARENTDISCOV YES
SELECANI NO
SNPINFOFLAGREDUCINGFORMAT SNPINFO_teste_UNSORTED.txt
REDUCEGENOFILE TESTE_UNSORTED.txt
FILEANIMALSCHIPFIMPUTE scen_1.txt
NCLUSTERS_K 3
SNPLISTTOPRINT snpListToDebug.txt
// ================== IMPUTATION ACCURACY ======================================
FOLDERLOCATIONORIGFILE /home/ricardo/workspace/0_RTBM_1001
GENO(FIMPUTE)FILE genotype_FIMPUTEFORMAT_BEFORE.txt
FOLDERLOCATIONIMPUTEDFILE /home/ricardo/workspace/0_RTBM_1001
GENO(IMPUTED)FILE genotype_FIMPUTEFORMAT_AFTER.txt
MAPFIMPUTEHD_BEFORE snp_infoFIMPUTE.txt
MAPFIMPUTEHD_AFTER snp_infoFIMPUTEResult.txt
ANIMALLISTTOCHECKACC animalListAcc.txt
// (PLINK/REDUCE/REDUCE_U/FIMPUTECHIP/PLOTR/FINDSNP/IMPACC/EXPAND/SORT/MAF/MAFFI) ==
// (HAM_LONG) =================================================================
SPECIFICFUNCTION PLOTR
// ##############################################################################
```


[Header]
GSGT Version 1.9.4
Nao envie ID de animais utilizando espaços como parte do nome …
ARS-BFGL-BAC-10172 X439Y G G G G B B 0.9506 0.027 1.000
ARS-BFGL-BAC-1020 X439Y - - - - - - 0.0000 0.177 0.630
ARS-BFGL-BAC-10245 X439Y C C G G B B 0.7579 0.160 1.424
ARS-BFGL-BAC-10345 X439Y A C A C A B 0.9276 0.679 0.870
ARS-BFGL-BAC-10365 X439Y G G C C B B 0.5335 0.076 1.188
ARS-BFGL-BAC-10375 X439Y A G A G A B 0.9567 0.372 0.418
ARS-BFGL-BAC-10591 X439Y A G A G A B 0.9003 0.311 0.481
ARS-BFGL-BAC-10867 X439Y G C C G A B 0.9434 0.428 0.468
ARS-BFGL-BAC-10919 X439Y A G A G A B 0.8526 0.876 0.982
ARS-BFGL-BAC-10951 X439Y T T A A A A 0.5140 0.601 0.034
ARS-BFGL-BAC-10952 X439Y A A A A A A 0.9512 1.331 0.033
ARS-BFGL-BAC-10960 X439Y - - - - - - 0.0000 0.047 0.241
ARS-BFGL-BAC-10972 X439Y G C C G A B 0.8759 1.016 0.877
ARS-BFGL-BAC-10975 X439Y A G A G A B 0.7379 0.525 0.443
ARS-BFGL-BAC-10986 X439Y G G C C B B 0.9309 0.045 0.877
ARS-BFGL-BAC-10993 X439Y T C A G A B 0.9014 1.087 0.986
ARS-BFGL-BAC-11000 X439Y T T A A A A 0.9686 0.578 0.023
ARS-BFGL-BAC-11003 X439Y T T A A A A 0.9215 1.576 0.060
ARS-BFGL-BAC-11007 X439Y T C A G A B 0.9454 0.689 0.695
ARS-BFGL-BAC-11025 X439Y - - - - - - 0.0000 0.171 0.808
ARS-BFGL-BAC-11028 X439Y A A A A A A 0.9678 0.427 0.019
ARS-BFGL-BAC-11034 X439Y T C A G A B 0.9509 0.591 0.696
ARS-BFGL-BAC-11039 X439Y C C G G B B 0.9658 0.017 0.982
ARS-BFGL-BAC-11042 X439Y G G G G B B 0.8506 0.019 1.305
ARS-BFGL-BAC-11044 X439Y C C G G B B 0.9654 0.022 1.332
ARS-BFGL-BAC-11047 X439Y T T A A A A 0.9465 0.983 0.020
ARS-BFGL-BAC-11050 X439Y T T A A A A 0.9212 1.009 0.047
ARS-BFGL-BAC-11057 X439Y G G G G B B 0.9023 0.101 0.857
ARS-BFGL-BAC-11061 X439Y - - - - - - 0.0000 0.179 0.933
ARS-BFGL-BAC-11090 X439Y G G G G B B 0.9417 0.076 0.809
ARS-BFGL-BAC-11101 X439Y A A A A A A 0.8870 0.699 0.022
Cuidado com os diferentes tipos de orientação !!!!
ARS-BFGL-BAC-11193 X439Y A A A A A A 0.9748 1.201 0.008
ARS-BFGL-BAC-11200 X439Y T T A A A A 0.6973 0.373 0.035
ARS-BFGL-BAC-11206 X439Y G G G G B B 0.9139 0.092 1.037
ARS-BFGL-BAC-11215 X439Y A A A A A A 0.4502 0.327 0.058
ARS-BFGL-BAC-11218 X439Y A A A A A A 0.9363 1.185 0.026
ARS-BFGL-BAC-11276 X439Y A A A A A A 0.9426 1.174 0.027
ARS-BFGL-BAC-11281 X439Y G G G G B B 0.8449 0.120 1.251
ARS-BFGL-BAC-11283 X439Y A A A A A A 0.8673 2.176 0.011
ARS-BFGL-BAC-11284 X439Y A A A A A A 0.9036 1.206 0.150
ARS-BFGL-BAC-11290 X439Y G G C C B B 0.8221 0.093 1.673

res.illumina.com/documents/products/technotes/technote_topbot.pdf

Illumina® SNP Genotyping

TECHNICAL NOTE

"TOP/BOT" Strand and "A/B" Allele

A guide to Illumina's method for determining Strand and Allele for the GoldenGate® and Infinium™ Assays.

INTRODUCTION

To address DNA strand designation and orientation for both human and non-human species, Illumina has developed a consistent and simple method to ensure uniformity in the reporting of genotype calls.

The method that Illumina has developed uses the top (TOP) and bottom (BOT) designations based on the polymorphism itself, or the contextual surrounding sequence. This document provides a description of the TOP/BOT method, as well as the generalized nomenclature of "Allele A" and "Allele B" within the Illumina genotyping system. Beginning in mid-2005, dbSNP also adopted this TOP/BOT nomenclature and has included consistently designate the same SNP orientation and allele calls even if public SNP databases and genome assemblies change. This will enable researchers worldwide to easily correlate the genotype calls made today to research that may have been completed several years ago. Researchers can be confident that the same genotype calls are being made across time.

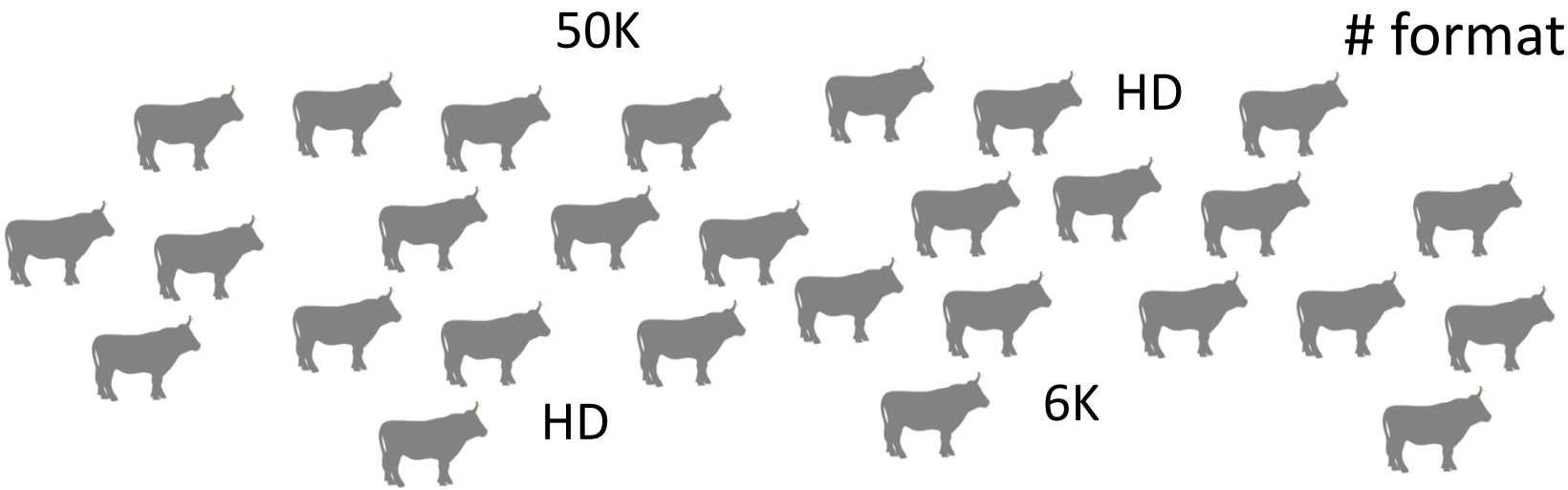
Additionally, although the human genome has been annotated and extensively SNP genotyped, the study of non-human species is growing rapidly. Much of this work is being done on species for which SNP databases are in their initial stages or do not yet exist. Many researchers that study both humans and non-human species also rely

Combinando sequência e genótipos

Forward Forward

Auxiliar construção dos haplótipos para imputação

777K para Forward Forward e não Forward Forward para AB



Sincronização / Backup: Solução não otimizada

Rotinas / Código fonte

rsync –avz --delete ../workspace rventura@uoguelph.ca:/home

Genótipos

rsync –avz --delete ../genotypes rventura@uoguelph.ca:/home

rsync –avz --delete ../genotypes rventura@server1:/home

rsync –avz --delete ../genotypes rventura@server2:/home

Sequência

rsync –avz --delete ../genotypes E:/externalharddrive

rsync –avz --delete ../genotypes E:/externalharddrive

Dados ordenados por Chr/Pos …

```
Name Chr Pos
Hapmap43437-BTA-101873  1       135098
ARS-BFGL-NGS-16466      1       267940
ARS-BFGL-NGS-19289      1       305793
ARS-BFGL-NGS-105096     1       353745
Hapmap34944-BES1_Contig627_1906 1       393248
BTA-07251-no-rs 1       434180
ARS-BFGL-NGS-98142      1       471078
Hapmap53946-rs29015852  1       516404
ARS-BFGL-NGS-114208     1       533815
ARS-BFGL-NGS-66449      1       571340
ARS-BFGL-NGS-51647      1       654413
ARS-BFGL-BAC-32770      1       845494
ARS-BFGL-NGS-65067      1       883895
ARS-BFGL-BAC-31497      1       905632
ARS-BFGL-BAC-32722      1       929617
ARS-BFGL-BAC-34682      1       950841
ARS-BFGL-NGS-3964       1       974586
ARS-BFGL-NGS-98203      1       1009504
ARS-BFGL-BAC-36895      1       1078256
```

```
ID CALL ...
CHAR8                    10250502525012111
```

E se nao estiver ordenado? ☺

Dimensão de sequência?

Outras funcionalidades …

```
// ##########################################################################
// ================ MANDATORY INFORMATION ===================================
GENOFILE_0125 genotype_9.txt
NPARALELJOBS 3
MAP(NCP) snp_info_54609.txt
FOLDERLOCATION /home/ricardo/workspace/0_RTBM_1001
FLAGWINLINUX /
// ============== QUALITY CONTROL (NO CR_SNP YET) ==========================
CR_ANI NO
MAF NO
CR_SNP NO
// ================== GENERAL FUNCTIONS =====================================
FLAGPARENTDISCOV YES
SELECANI NO
SNPINFOFLAGREDUCINGFORMAT SNPINFO_teste_UNSORTED.txt
REDUCEGENOFILE TESTE_UNSORTED.txt
FILEANIMALSCHIPFIMPUTE scen_1.txt
NCLUSTERS_K 3
SNPLISTTOPRINT snpListToDebug.txt
// ================== IMPUTATION ACCURACY ===================================
FOLDERLOCATIONORIGFILE /home/ricardo/workspace/0_RTBM_1001
GENO(FIMPUTE)FILE genotype_FIMPUTEFORMAT_BEFORE.txt
FOLDERLOCATIONIMPUTEDFILE /home/ricardo/workspace/0_RTBM_1001
GENO(IMPUTED)FILE genotype_FIMPUTEFORMAT_AFTER.txt
MAPFIMPUTEHD_BEFORE snp_infoFIMPUTE.txt
MAPFIMPUTEHD_AFTER snp_infoFIMPUTEResult.txt
ANIMALLISTTOCHECKACC animalListAcc.txt
// (PLINK/REDUCE/REDUCE_U/FIMPUTECHIP/PLOTR/FINDSNP/IMPACC/EXPAND/SORT/MAF/MAFFI) ==
// (HAM_LONG) ==============================================================
SPECIFICFUNCTION PLOTR
// ##########################################################################
```

Codificando arquivos de diferentes laboratórios

B1310001A100A

Fonte dos dados: BIO, Colaborador ou LLOC

Ano Genotipagem

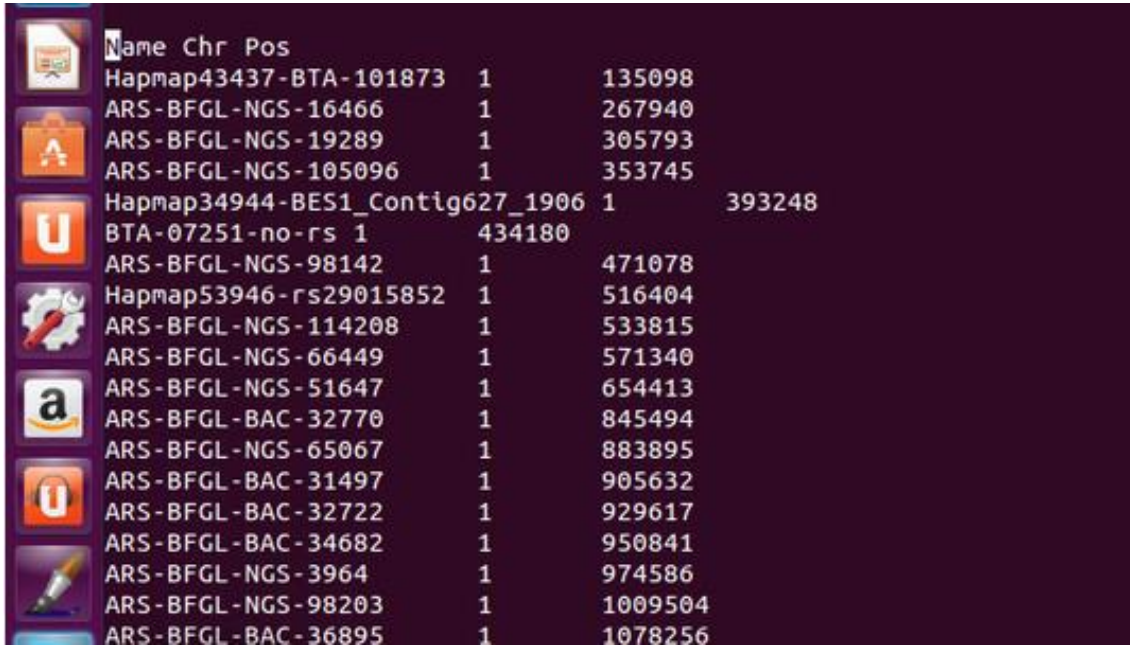
Numero sequencial de genotipagem para controle interno

Densidade

Lote + Laboratorio

	A	B	C	D	E	F	G	H	I	J	K
1	B	13	10001	A	100	A	B1310001A100A	Number of Samples	50		
2	B	13	10002	A	100	A	B1310002A100A	Initial Value	10001		
3	B	13	10003	A	100	A	B1310003A100A	Source (B,C,LLOC)	B		
4	B	13	10004	A	100	A	B1310004A100A	Year	13		
5	B	13	10005	A	100	A	B1310005A100A	Lab	A		
6	B	13	10006	A	100	A	B1310006A100A	Shipment	100		
7	B	13	10007	A	100	A	B1310007A100A	Density	A		
8	B	13	10008	A	100	A	B1310008A100A				
9	B	13	10009	A	100	A	B1310009A100A				
10	B	13	10010	A	100	A	B1310010A100A	Densities:		Lab:	
11	B	13	10011	A	100	A	B1310011A100A	A	50K	A	Delta - UoA
12	B	13	10012	A	100	A	B1310012A100A	B	6K	B	Geneseek
13	B	13	10013	A	100	A	B1310013A100A	C	777,962 Illumina	C	Deoxi Brazil
14	B	13	10014	A	100	A	B1310014A100A	D	648,484 Affy	D	DNA Landma
15	B	13	10015	A	100	A	B1310015A100A	E	6KGGP		
16	B	13	10016	A	100	A	B1310016A100A	Source			
17	B	13	10017	A	100	A	B1310017A100A	B	BIO		
18	B	13	10018	A	100	A	B1310018A100A	C	Collaborator		
19	B	13	10019	A	100	A	B1310019A100A	L	LLOC		

Ivan Campos (IC): MySQL x Postgresql



```
Name Chr Pos
Hapmap43437-BTA-101873  1       135098
ARS-BFGL-NGS-16466      1       267940
ARS-BFGL-NGS-19289      1       305793
ARS-BFGL-NGS-105096     1       353745
Hapmap34944-BES1_Contig627_1906 1       393248
BTA-07251-no-rs 1       434180
ARS-BFGL-NGS-98142      1       471078
Hapmap53946-rs29015852  1       516404
ARS-BFGL-NGS-114208     1       533815
ARS-BFGL-NGS-66449      1       571340
ARS-BFGL-NGS-51647      1       654413
ARS-BFGL-BAC-32770      1       845494
ARS-BFGL-NGS-65067      1       883895
ARS-BFGL-BAC-31497      1       905632
ARS-BFGL-BAC-32722      1       929617
ARS-BFGL-BAC-34682      1       950841
ARS-BFGL-NGS-3964       1       974586
ARS-BFGL-NGS-98203      1       1009504
ARS-BFGL-BAC-36895      1       1078256
```

```
ID CALL ...
CHAR8                    10250502525012111
LIMO9                    01250520525022110
```

Banco de dados

ID Tatoo
Genotype_6K
Genotype_9K
Genotype_50K
Genotype_777
Genotype_648

ID Tatoo
Genotype_50K_Imputed
Genotype_777_Imputed
Genotype_648_Imputed

ID Tatoo
DataGenotipagem
Laboratorio
CallRate
TipoAmostra
ContatoLaboratorio
Remessa
Densidade
Orientação

Armazenar baseado no Animal ou SNP?

Banco de dados

RTBM, Ireland files, desenho do banco de dados, rsync, etc e tal?

SNPchiMp V.2

bioinformatics.tecnoparco.org/SNPchimp/
Esta página está em inglês Deseja traduzi-la? Traduzir Não Nunca traduzir do inglês
SNPchiMp v.2
A multi-species database to disentangle the SNP chip jungle
HOME INFO DOWNLOAD BROWSE DATA SOURCE CONTACTS FAQS NEWS LINKS LOGIN
GENE 2 FARM
SEVENTH FRAMEWORK PROGRAMME
SNPchiMp v.2
A multi-species database to disentangle the SNP chip jungle

SNPchiMp V.2

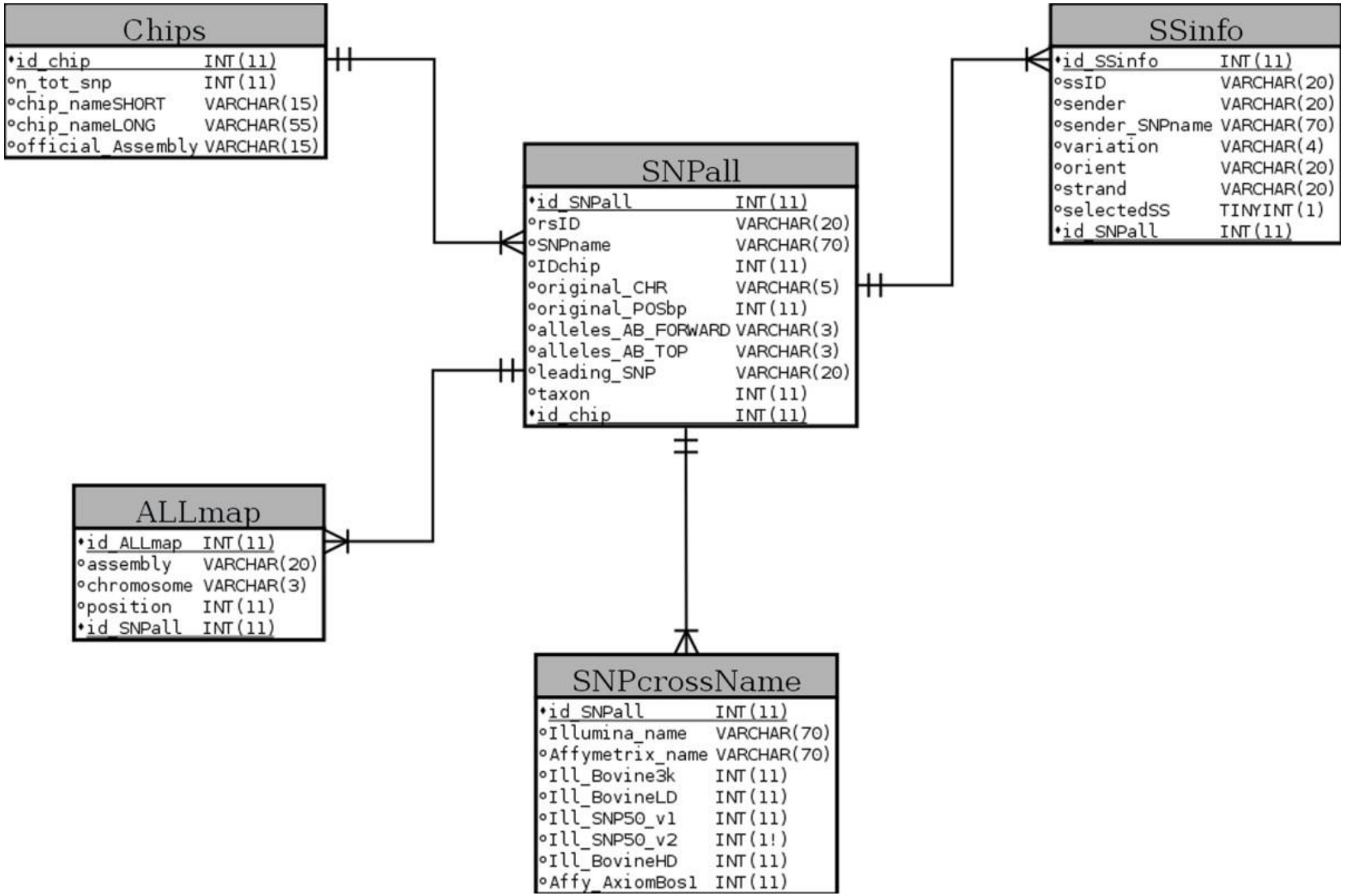
Atualizar posição de qualquer SNP em qualquer versão ASSEMBLY

Informações de orientação, strand para futuras integrações dos genotipos com dados de sequência

Acesso a diferentes codificações alélicas da Illumina :
FORWARD REVERSE
TOP BOT

Navegação direta dos SNPs resultantes para NCBI e ENSEMBL

Banco de dados



SNPchiMp v.2

A multi-species database to disentangle the SNP chip jungle

HOME | INFO | DOWNLOAD | BROWSE | DATA SOURCE | CONTACTS | FAQS | NEWS | LINKS | LOGIN

You are here: Home ▸ Download ▸ Download Cow Data

Chosen Species: Cow

Step 1: Please select the SNP chip information desired:

- [ ] Illumina Bovine3k BeadChip (2,900 SNPs)
- [ ] Illumina BovineLD BeadChip (6,909 SNPs)
- [ ] Illumina BovineLD v1.1 BeadChip (6912 SNPs)
- [ ] Illumina BovineSNP50v1 BeadChip (54,001 SNPs)
- [x] Illumina BovineSNP50v2 BeadChip (54,609 SNPs)
- [ ] Illumina BovineHD BeadChip (777,962 SNPs)

- [ ] GeenSeek Genomic Profiler LD v1 (8,610 SNPs)
- [ ] GeneSeek Genomic Profiler LD v2 (19,721 SNPs)
- [ ] GeneSeek Genomic Profiler LD v3 (26,151 SNPs)
- [ ] GeneSeek Genomic Profiler HD (76,879 SNPs)

- [ ] Affymetrix Axiom ® Bovine (648,875 SNP probes)

Step 3: Please select which information you want to display:

Assembly:

- (•) Native platform (Source: producer)
- ( ) UMD 3.1 (Source: dbSNP)
- ( ) BTAU 4.2 (Source: dbSNP)
- ( ) BTAU 4.6 (Source: dbSNP)

Chromosome and Position all

- [ ] *ss information*

- [ ] *Exchange Interbull Index*

Allele coding:

- [x] A/B forward alleles (Illumina Only)
- [x] A/B top alleles (Illumina Only)

SNPchiMp V.2

chip_name	rs	Alleles_A_B_FORWARD	Alleles_A_B_TOP	chromosome	position	SNP_name
50Kv2	rs110605804	T/C	A/G	0	0	ARS-BFGL-BAC-11602
50Kv2	rs43710845	A/G	A/G	0	0	ARS-BFGL-BAC-13205
50Kv2	rs41766931	A/G	A/G	0	0	ARS-BFGL-BAC-19353
50Kv2	rs109133222	T/G	A/C	0	0	ARS-BFGL-BAC-27150
50Kv2	rs109534145	A/C	A/C	0	0	ARS-BFGL-BAC-27427
50Kv2	rs110676719	A/C	A/C	0	0	ARS-BFGL-BAC-3012
50Kv2	rs110662702	T/G	A/C	0	0	ARS-BFGL-BAC-31012
50Kv2	rs109817599	T/C	A/G	0	0	ARS-BFGL-BAC-33662
50Kv2	NULL	T/C	A/G	0	0	ARS-BFGL-BAC-35028
50Kv2	rs110843183	T/C	A/G	0	0	ARS-BFGL-BAC-3582
50Kv2	rs81109693	T/C	A/G	0	0	ARS-BFGL-BAC-35948
50Kv2	rs109754744	T/C	A/G	0	0	ARS-BFGL-NGS-100012
50Kv2	rs42588436	T/C	A/G	0	0	ARS-BFGL-NGS-100232

SNPchiMp V.2

HOME | INFO | DOWNLOAD | BROWSE | DATA SOURCE

You are here: Home ▸ Download ▸ Download Cow Data

Chosen Species: Cow

Step 1: Please select the SNP chip information desired:

- [ ] Illumina Bovine3k BeadChip (2,900 SNPs)
- [ ] Illumina BovineLD BeadChip (6,909 SNPs)
- [ ] Illumina BovineLD v1.1 BeadChip (6912 SNPs)
- [ ] Illumina BovineSNP50v1 BeadChip (54,001 SNPs)
- [x] Illumina BovineSNP50v2 BeadChip (54,609 SNPs)
- [x] Illumina BovineHD BeadChip (777,962 SNPs)

- [ ] GeenSeek Genomic Profiler LD v1 (8,610 SNPs)
- [ ] GeneSeek Genomic Profiler LD v2 (19,721 SNPs)
- [ ] GeneSeek Genomic Profiler LD v3 (26,151 SNPs)
- [ ] GeneSeek Genomic Profiler HD (76,879 SNPs)

- [ ] Affymetrix Axiom ® Bovine (648,875 SNP probes)

Step 2: Type of information required:

(Commercial SNP ID and rs ID are displayed by default)

- ( ) Detailed SNP information
- (•) Across SNPchip Table

SNPchiMp V.2

SNPchimp_result_1861101345 - Bloco de notas

Arquivo Editar Formatar Exibir Ajuda

```
AX-20847703 BovineHD15000240850 1       rs43764409      BovineHD21000091330     1       rs43764424      BovineHD03000262070
s43764762       BovineHD09000010360     1       rs43764785      BovineHD10000259710     1       rs43764803      BovineHD110
neHD01000252660 1       rs43765070      BovineHD01000252670     1       rs43765077      BovineHD01000152720     1       rs4
410     BovineHD16000237450     1       rs43765431      BovineHD03000292750     1       rs43765447      BovineHD06000254190
4       BovineHD09000092880     1       rs43765777      AX-27798677 BovineHD09000092920 1       rs43765781      BovineHD090
rs43766134      BovineHD16000207960     1       rs43766161      BovineHD16000133150     1       rs43766181      BovineHD160
2780    1       rs43766558      BovineHD04000122790     1       rs43766560      BovineHD04000122800     1       rs43766571
eHD04000135250  1       rs43766835      BovineHD41000018610     1       rs43766836      BovineHD07000091860     1       rs4
D29000079380    1       rs43767149      BovineHD26000063590     1       rs43767150      BovineHD26000063600     1       rs4
BovineHD03000353310     1       rs43767455      BovineHD11000169150     1       rs43767478      BovineHD09000092910     1
s43767816       BovineHD03000191570     1       rs43767817      BovineHD03000232720     1       rs43767845      BovineHD010
1401331 BovineHD17000168350     1       rs43768221      BovineHD17000185890     1       rs43768264      BovineHD03000262000
656     BovineHD12000142840     1       rs43768665      BovineHD07000091890     1       rs43768674      BovineHD07000237240
1       BovineHD06000043730     1       rs43768937      BovineHD08000095890     1       rs43768944      BovineHD08000096990
000154480       1       rs43769253      BovineHD15000073620     1       rs43769257      BovineHD11000070710     1       rs4
0321510 1       rs43769699      BovineHD15000061630     1       rs43769701      BovineHD15000063760     1       rs43769716
neHD04000069590 1       rs43770125      BovineHD02000385410     1       rs43770141      BovineHD15000015430     1       rs4
506     AX-20980650 BovineHD16000090150 1       rs43770513      BovineHD01000088570     1       rs43770567      BovineHD160
0       1       rs43770846      BovineHD13000157320     1       rs43770882      BovineHD14000007560     1       rs43770889
6       BovineHD18000055030     1       rs43771169      BovineHD18000055050     1       rs43771172      BovineHD18000055060
771541  BovineHD07000250760     1       rs43771542      BovineHD07000250860     1       rs43771548      BovineHD08000167950
2210    BovineHD24000084540     1       rs43772212      BovineHD24000084530     1       rs43772247      BovineHD17000032830
4710    1       rs43772685      BovineHD08000168000     1       rs43772701      BovineHD07000250770     1       rs43772714
43773050        BovineHD02000024360     1       rs43773060      BovineHD21000038820     1       rs43773061      BovineHD210
1       rs43773468      AX-25046039 BTB-014362900       1       rs43773489      BovineHD02000024501     1       rs43773490
rs55617213      BovineHD30000408031     0       rs55617218      UA-IFASA-76160  1       rs55617231      BovineHD14000024761
s81109482       BTA-110823-no-rs0       1       rs81109494      BovineHD01000203901     1       rs81109506      BTA-23530-n
map25980-BTC-0170341    0       rs81122201      Hapmap22899-BTC-0180011 1       rs81123300      Hapmap31085-BTA-1292841 1
apmap27583-BTA-1351601  1       rs81134715      Hapmap24661-BTA-1354521 1       rs81135172      Hapmap29024-BTA-1357121 1
Hapmap23854-BTC-0624121 1       rs81154068      Hapmap24324-BTC-0624491 1       rs81154138      Hapmap26150-BTA-1462751 1
6       Hapmap24292-BTC-0709831 1       rs81165579      Hapmap31151-BTA-1525891 1       rs81166347      Hapmap31476-BTA-153
916     Hapmap28106-BTA-1579731 1       rs81172105      Hapmap32371-BTA-1581851 0       rs81172151      Hapmap24347-BTA-158
1       rs81182500      BovineHD19000094270     1       rs81183038      BovineHD22000117490     1       rs81183101      Bov
```

SNPchiMp V.2

	A	B	C	
783104	50Kv2	IlluHD	rs	SNP_name
783105	1	1	rs81177673	Hapmap31526-BTA-164250
783106	1	1	rs81177829	Hapmap23101-BTA-164419
783107	1	1	rs81177938	INRA-185
783108	1	1	rs81177940	AX-20126306 BPI-1
783109	1	1	rs81177941	AX-20126304 BPI-2
783110	1	1	rs81177946	UA-IFASA-6924
783111	1	1	rs81177947	UA-IFASA-7656
783112	0	1	rs81177949	UA-IFASA-5520
783113	1	1	rs81177950	UA-IFASA-7534
783114	0	1	rs81178564	BovineHD0200031895
783115	0	1	rs81179104	BovineHD0400026528
783116	0	1	rs81179743	BovineHD0600027756
783117	0	1	rs81180709	AX-18987119 BovineHD1000026789
783118	0	1	rs81181939	BovineHD1600017454
783119	0	1	rs81182500	BovineHD1900009427
783120	0	1	rs81183038	BovineHD2200011749
783121	0	1	rs81183101	BovineHD2200017533
783122	0	1	rs81184743	BovineHD3200000033
783123	0	1	rs81184754	BovineHD3200000072
783124	0	1	rs81184776	BovineHD3200000368
783125	0	1	rs81184782	BovineHD3200000384
783126	0	1	rs81184797	BovineHD3200000401
783127	0	1	rs8193046	AX-27654544 BovineHD0800032700

bioinformatics.tecnoparco.org/SNPchimp/index.php/browse

SNPchiMp v.2

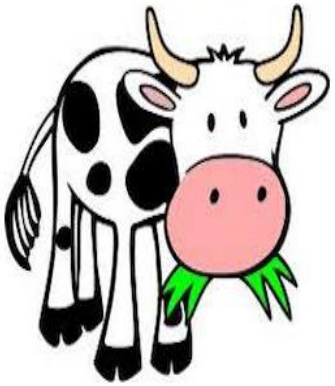
A multi-species database to disentangle the SNP chip jungle

HOME | INFO | DOWNLOAD | BROWSE | DATA SOURCE | CONTACTS | FAQS | NEWS | LINKS | LOGIN

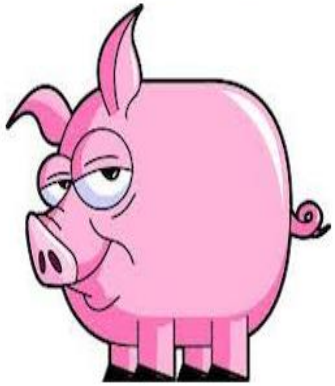
You are here: Home ▸ Browse

Please choose the desired species:

BOVINE



PORCINE



EQUINE

OVINE

CAPRINE

SNPchiMp v.2

A multi-species database to disentangle the SNP chip jungle

HOME | INFO | DOWNLOAD | BROWSE | DATA SOURCE | CONTACTS | FAQS | NEWS | LINKS | LOGIN

You are here: Home ▸ Browse ▸ Browse Cow Data

Chosen Species: Cow

Step 1: Please select one assembly:

- Native platform (Source: producer)
- UMD 3.1 (Source: dbSNP) (Biomart link available for this assembly)
- BTAU 4.2 (Source: dbSNP)
- BTAU 4.6 (Source: dbSNP)

Step 2: Information required:

(Commercial SNP ID and rs ID are displayed by default)

- [x] Present in Chip
- [ ] ss information

1:1..4000

Step 3: Write your query:

Browse SNPs by position — Syntax chromosome:start_position..stop_position

Browse SNPs by rs IDs — Syntax rsId1,rsId2,...,rsIdn

Browse SNPs by ss IDs — Syntax ssId1,ssId2,...,ssIdn

Browse SNPs by name — Syntax SNPname1,SNPname2,...,SNPnamen

SNPchiMp v.2

A multi-species database to disentangle the SNP chip jungle

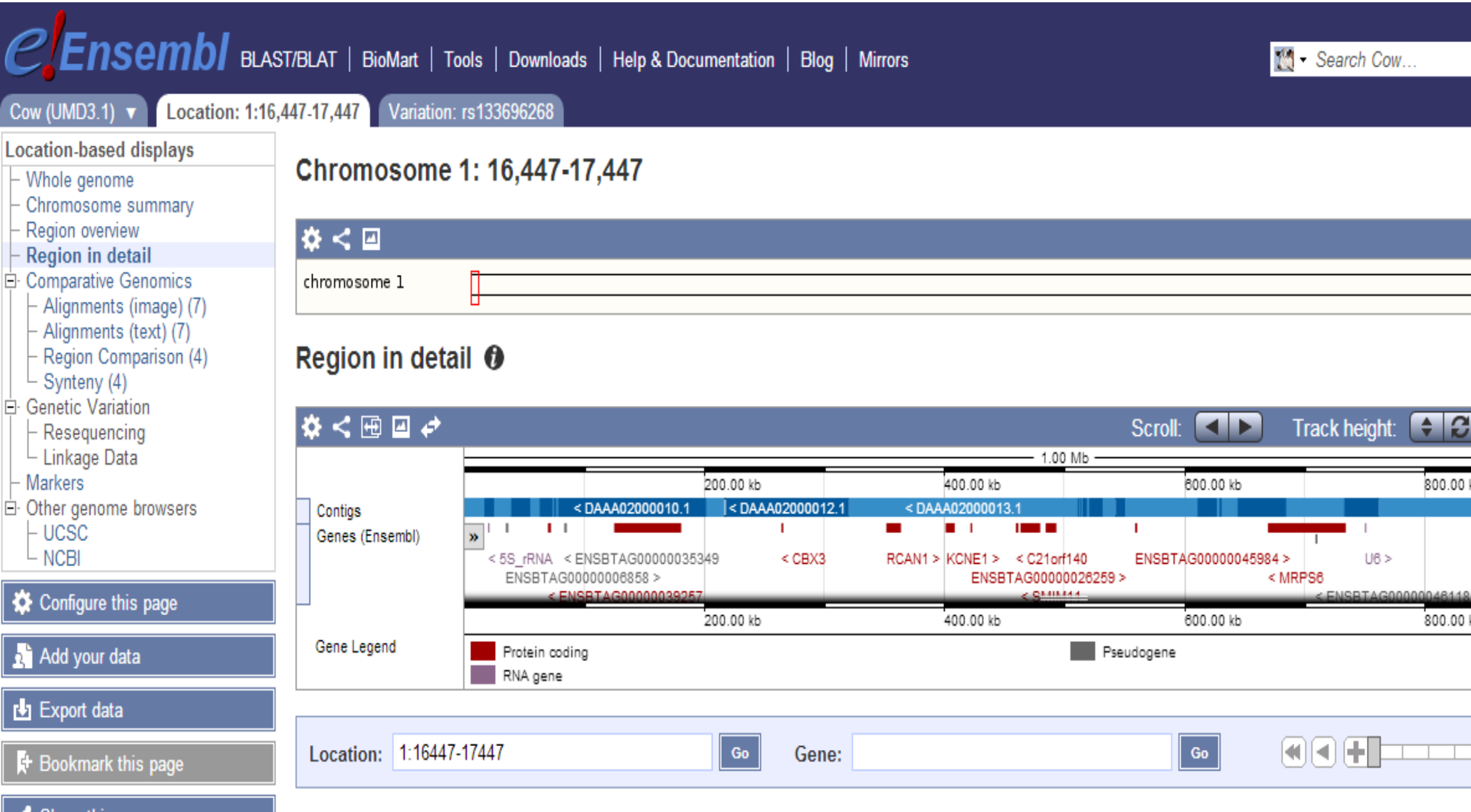
Query your results on *Ensembl BioMart data mining tool* (beta)

Download as Comma Separated file Tab Separated file

The result (up to 20 rows)

rs	chip_name	ss	alleles	orient	strand	sender	SNPname_sender	chromosome	position	SNP_name	ensembl	ncbi
rs133696268	IlluHD	ss250608777	A/C	forward	top	BFGL_BARC_USDA	BovineHD0100000005	1	16947	BovineHD0100000005	Ensembl	NCB
rs132895573	IlluHD	ss250608779	A/G	forward	top	BFGL_BARC_USDA	BovineHD0100000015	1	36337	BovineHD0100000015	Ensembl	NCB
rs133696268	AffyHD	NULL	NULL	NULL	NULL	NULL	NULL	1	16947	AX-18000040	Ensembl	NCB
rs208669904	AffyHD	NULL	NULL	NULL	NULL	NULL	NULL	1	36440	AX-18000110	Ensembl	NCB

SNPchiMp V.2


e!Ensembl
BLAST/BLAT | BioMart | Tools | Downloads | Help & Documentation | Blog | Mirrors
Search Cow...
Cow (UMD3.1)
Location: 1:16,447-17,447
Variation: rs133696268
Location-based displays
Whole genome
Chromosome summary
Region overview
Region in detail
Comparative Genomics
Alignments (image) (7)
Alignments (text) (7)
Region Comparison (4)
Synteny (4)
Genetic Variation
Resequencing
Linkage Data
Markers
Other genome browsers
UCSC
NCBI
Configure this page
Add your data
Export data
Bookmark this page
Chromosome 1: 16,447-17,447
chromosome 1
Region in detail
Scroll:
Track height:
1.00 Mb
200.00 kb
400.00 kb
600.00 kb
800.00 kb
Contigs
Genes (Ensembl)
< DAAA02000010.1
< DAAA02000012.1
< DAAA02000013.1
< 5S_rRNA
< ENSBTAG00000035349
ENSBTAG00000006858 >
< CBX3
RCAN1 >
KCNE1 >
< C21orf140
ENSBTAG00000026259 >
ENSBTAG00000045984 >
< MRPS6
U6 >
Gene Legend
Protein coding
RNA gene
Pseudogene
Location:
1:16447-17447
Go
Gene:
Go

SNPchiMp v.2

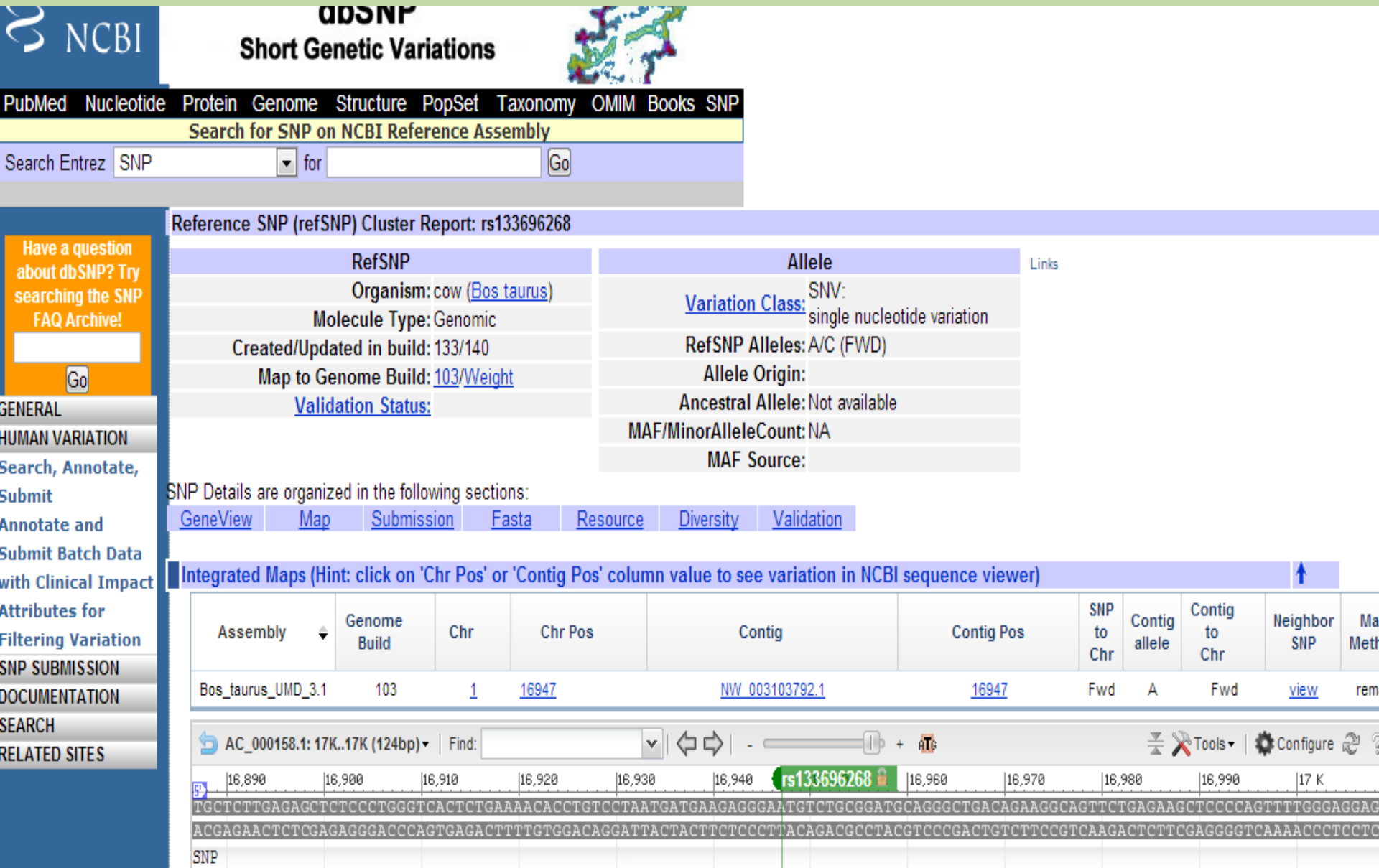
A multi-species database to disentangle the SNP chip jungle

Query your results on *Ensembl BioMart data mining tool* (beta)

Download as Comma Separated file Tab Separated file

The result (up to 20 rows)

rs	chip_name	ss	alleles	orient	strand	sender	SNPname_sender	chromosome	position	SNP_name	ensembl	ncbi
rs133696268	IlluHD	ss250608777	A/C	forward	top	BFGL_BARC_USDA	BovineHD0100000005	1	16947	BovineHD0100000005	Ensembl	NCB
rs132895573	IlluHD	ss250608779	A/G	forward	top	BFGL_BARC_USDA	BovineHD0100000015	1	36337	BovineHD0100000015	Ensembl	NCB
rs133696268	AffyHD	NULL	NULL	NULL	NULL	NULL	NULL	1	16947	AX-18000040	Ensembl	NCB
rs208669904	AffyHD	NULL	NULL	NULL	NULL	NULL	NULL	1	36440	AX-18000110	Ensembl	NCB


NCBI
dbSNP
Short Genetic Variations
PubMed Nucleotide Protein Genome Structure PopSet Taxonomy OMIM Books SNP
Search for SNP on NCBI Reference Assembly
Search Entrez SNP for Go
Have a question about dbSNP? Try searching the SNP FAQ Archive!
Go
GENERAL
HUMAN VARIATION
Search, Annotate, Submit
Annotate and Submit Batch Data with Clinical Impact Attributes for Filtering Variation
SNP SUBMISSION
DOCUMENTATION
SEARCH
RELATED SITES
Reference SNP (refSNP) Cluster Report: rs133696268
RefSNP
Organism: cow (Bos taurus)
Molecule Type: Genomic
Created/Updated in build: 133/140
Map to Genome Build: 103/Weight
Validation Status:
Allele
Links
Variation Class: SNV: single nucleotide variation
RefSNP Alleles: A/C (FWD)
Allele Origin:
Ancestral Allele: Not available
MAF/MinorAlleleCount: NA
MAF Source:
SNP Details are organized in the following sections:
GeneView Map Submission Fasta Resource Diversity Validation
Integrated Maps (Hint: click on 'Chr Pos' or 'Contig Pos' column value to see variation in NCBI sequence viewer)
Assembly
Genome Build
Chr
Chr Pos
Contig
Contig Pos
SNP to Chr
Contig allele
Contig to Chr
Neighbor SNP
Bos_taurus_UMD_3.1 103 1 16947 NW_003103792.1 16947 Fwd A Fwd view
AC_000158.1: 17K..17K (124bp)
Find:
Tools
Configure
16,890 16,900 16,910 16,920 16,930 16,940 rs133696268 16,960 16,970 16,980 16,990 17 K
TGCTCTTGAGAGCTCTCCCTGGGTCACTCTGAAAACACCTGTCCTAATGATGAAGAGGGAATGTCTGCGGATGCAGGGCTGACAGAAGGCAGTTCTGAGAAGCTCCCCAGTTTTGGGAGGAG
ACGAGAACTCTCGAGAGGGACCCAGTGAGACTTTTGTGGACAGGATTACTACTTCTCCCTTACAGACGCCTACGTCCCGACTGTCTTCCGTCAAGACTCTTCGAGGGGTCAAAACCCTCCTC
SNP

SNPchiMp v.2

A multi-species database to disentangle the SNP chip jungle

HOME | INFO | DOWNLOAD | BROWSE | DATA SOURCE | CONTACTS | FAQS | NEWS | LINKS | LOGIN

You are here: Home ▸ Browse ▸ Browse Cow Data

Chosen Species: Cow

Step 1: Please select one assembly:

- ○ Native platform (Source: producer)
- ◉ UMD 3.1 (Source: dbSNP) (Biomart link available for this assembly)
- ○ BTAU 4.2 (Source: dbSNP)
- ○ BTAU 4.6 (Source: dbSNP)

Step 2: Information required:

(Commercial SNP ID and rs ID are displayed by default)

- ☑ Present in Chip
- ☑ ss information

Step 3: Write your query:

Browse SNPs by position — Syntax chromosome:start_position..stop_position

Browse SNPs by rs IDs — rs133696268

Browse SNPs by ss IDs — Syntax ssId1,ssId2,...,ssIdn

Browse SNPs by name — Syntax SNPname1,SNPname2,...,SNPnamen

SNPchiMp v.2

A multi-species database to disentangle the SNP chip jungle

Query your results on *Ensembl BioMart data mining tool* (beta)

Download as Comma Separated file Tab Separated file

The result (up to 20 rows)

rs	chip_name	ss	alleles	orient	strand	sender	SNPname_sender	chromosome	position	SNP_name	ensembl	ncbi
rs133696268	IlluHD	ss250608777	A/C	forward	top	BFGL_BARC_USDA	BovineHD0100000005	1	16947	BovineHD0100000005	Ensembl	NCBI
rs133696268	AffyHD	NULL	NULL	NULL	NULL	NULL	NULL	1	16947	AX-18000040	Ensembl	NCBI

SNPchiMp v.2

A multi-species database to disentangle the SNP chip jungle

Chosen Species: Cow

Step 1: Please select one assembly:

- ○ Native platform (Source: producer)
- ◉ UMD 3.1 (Source: dbSNP) (Biomart link available for this assembly)
- ○ BTAU 4.2 (Source: dbSNP)
- ○ BTAU 4.6 (Source: dbSNP)

Step 2: Information required:

(Commercial SNP ID and rs ID are displayed by default)

Step 3: Write your query:

Browse SNPs by position: Syntax chromosome:start_positio

Browse SNPs by rs IDs: Syntax rsId1,rsId2,...,rsIdn

Browse SNPs by ss IDs: Syntax ssId1,ssId2,...,ssIdn

Browse SNPs by name: ARS-BFGL-BAC-10591

SNPchiMp v.2

A multi-species database to disentangle the SNP chip jungle

HOME | INFO | DOWNLOAD | BROWSE | DATA SOURCE | CONTACTS | FAQS

Query your results on *Ensembl BioMart data mining tool* (beta)

Download as Comma Separated file Tab Separated file

The result (up to 20 rows)

rs	SNP_name	chromosome	position	ensembl	ncbi
rs109076417	ARS-BFGL-BAC-10591	14	17544926	e! Ensembl	NCBI

SNPchiMp V.2

The following are the data sources used to build the database.

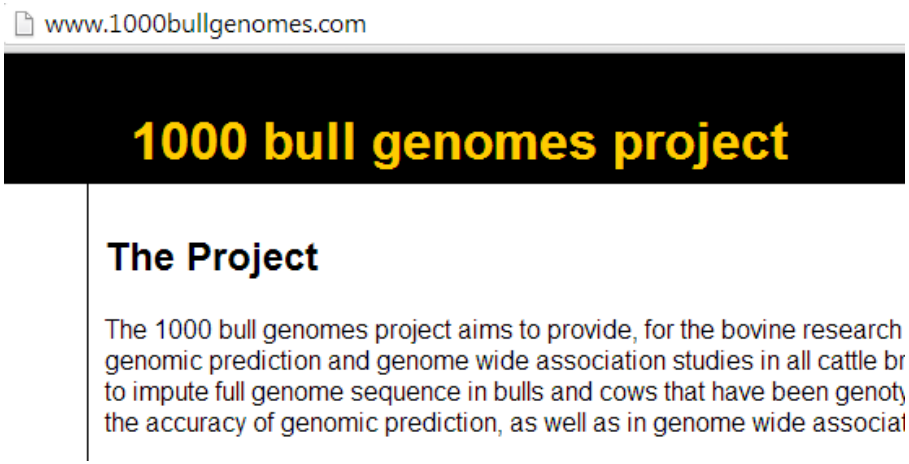
Sources of data: COW species

SNP chip	SNP names	Native positions	Allele coding	dbSNP builds	Interbull info
Illumina 3k	iManifest (Illumina)	iManifest (Illumina)	iManifest (Illumina)	136,137,138,140	Interbull
Illumina LD v1	iManifest (Illumina)	iManifest (Illumina)	iManifest (Illumina)	136,137,138,140	Interbull
Illumina LD v1.1	iManifest (Illumina)	iManifest (Illumina)	*iManifest (Illumina)*	*136,137,138,140*	N/A** NEW!!
Illumina 54k v.1	iManifest (Illumina)	iManifest (Illumina)	iManifest (Illumina)	136,137,138,140	Interbull
Illumina 54k v.2	iManifest (Illumina)	iManifest (Illumina)	iManifest (Illumina)	136,137,138,140	Interbull
Illumina HD	iManifest (Illumina)	iManifest (Illumina)	iManifest (Illumina)	136,137,138,140	Interbull
Affymetrix HD	Axiom GW Bos v.34	Axiom GW Bos v.34	Axiom GW Bos v.34	136,137,138,140	N/A
Geneseek GPP v.1	iManifest (Geneseek)	iManifest (Geneseek)	iManifest (Geneseek)	136,137,138,140	Interbull
Geneseek GPP v.2	SNP list (Geneseek)	linked to Illumina[1]	linked to Illumina[1]	136,137,138,140	Interbull
Geneseek GPP v.3	iManifest (Geneseek)	iManifest (Geneseek)	iManifest (Geneseek)	136,137,138,140	N/A** NEW!!
Geneseek GPP HD	SNP list (Geneseek)	linked to Illumina[1]	linked to Illumina[1]	136,137,138,140	Interbull

[1] *Illumina information corresponding to these SNP names were used to obtain such information (which might be incomplete)*

Colaboração internacional: compartilhar genótipos?

~ 200.000 animals genotipados e compartilhados



: 1000 animais sequenciados

Novas oportunidades para compartilhamento de genótipos