

GenoEx-PSE File Exchange Formats

12 August 2016

Authorized Services Users with a signed Service Agreement submitted to ICAR and the Interbull Centre will have access for downloading SNP genotypes and associated data to the GenoEx-PSE database and will have the obligation to routinely upload required data to the GenoEx-PSE database. The SNP genotypes and related information submitted (uploaded) to GenoEx-PSE is categorized in two groups:

- 1) Information related to a group of animals with a SNP genotype to upload;
- 2) Information related to the actual SNP genotypes for the animals in 1).

At the time of each upload from a Service User, both files must be submitted at the same time. Data submitted in one file will not be processed until both files are available.

Information Related to a Group of Animals

The following table describes the file format to be used for uploading general information for the group of animals for which SNP genotypes will be submitted in the associated second file (see below). This format will also be used when downloading data from GenoEx-PSE. This file will be exchanged as a variable length, comma delimited file in .csv format and include a single record for each animal for which the SNP genotype details will be included in the second file. For example, if the second file (File 704-AB or 704-TOP) includes SNP genotype results for 100 animals, this files will have 100 records.

Field Name	Column Size	Example	Description
Record Type	3	702	Numeric
Service User	20	INRA	Alphanumeric
Source Country of Animals	3	FRA	Alpha
Animal ID - Breed Code	3	BSW	Alpha
Animal ID - Nation Code	3	AUS	Alphanumeric (USA, 840, CAN, GBR, FRA, DEU, ...)
Animal ID - Sex Code	1	M	Alpha (M or F)
Animal ID - Registration	18	A1234567890	Alphanumeric
Genotyping Laboratory Identification	20	Weatherbys, GeneSeek etc.	Alpha
Sample ID	20	R1234567890	Alphanumeric
Scan Date	8	yyyymmdd	Numeric
Platform	2	1=Illumina 2=Affymetrix	SNP Chip Platform Code as: 1 = Illumina microarrays 2 = Illumina Truseq (GBS) 3 = Affymetrix Axiom® Microarray 4 = Affymetrix /Eureka 5 = KASPar 6 = Fluidigm 7 = Sequenom 8 = SNPLINE™ 9 = OpenArray® 10 = TaqMan
No. SNPs in Genotype	6	55647	Total number of SNPs in the animal's full genotype used to create the SNP record for GenoEx-PSE
Genotype Call Rate	4	99.9	Percent call rate of the animal's full genotype used to create the SNP record for GenoEx-PSE
Maximum Record Length	111		

Information Related to the SNP Genotypes of Animals

The following table describes the file format to be used for uploading the specific SNP genotypes to the GenoEx-PSE database. The same file format will be used for producing the data files to be downloaded from the GenoEx-PSE database. The Service User may select to upload and/or download SNP genotype data in either the "AB" or "TOP" allele designations, which determines the content of the first field, namely Record Type, in the following file format (i.e.: File 704-AB versus File 704-TOP, respectively). This file will be exchanged as a variable length, comma delimited file in .csv format and include a single record for each SNP included for each animal. For example, if the first file (File 702) includes 100 animals and each SNP genotype include the 200 SNPs recommended by ISAG for Parentage Verification then this second file will include a maximum of 20,000 records (100 animals x 200 SNPs each). In the event that any SNP was not "called" and the result is missing, then that SNP for that animal should not be included in the GenoEx-PSE genotype exchange file. For this reason, even if the first file (File 702) includes 100 animals, this second file will not necessarily include a total of 20,000 records.

Field Name	Column Size	Example	Description
Record Type	7	704-AB (or 704-TOP)	Alphanumeric
Animal ID - Breed Code	3	BSW	Alpha
Animal ID - Nation Code	3	AUS	Alphanumeric (USA, 840, CAN, GBR, FRA, DEU, ...)
Animal ID - Sex Code	1	M	Alpha (M or F)
Animal ID - Registration	18	A1234567890	Alphanumeric
SNP Name	70	ARS-BFGL-BAC-19454	Alphanumeric
Allele 1	1	A/B or A/C/G/T	Alpha
Allele 2	1	A/B or A/C/G/T	Alpha
Maximum Record Length	104		