

**File 706**

for the upload of genomic data in GenoEx-GDE database.

**706 format**

|  |  |  |
| --- | --- | --- |
| **Field Description** | **Format** | **Example** |
| Record type 1 | integer 3 | 706 |

|  |  |  |
| --- | --- | --- |
| Breed of animal 4 | character 3 | BSW |
| Country of first registration of animal 2 | character 3 | AUS |

|  |  |  |
| --- | --- | --- |
| Sex | character 1 | M |
| ID number of animal 5 | alphanumeric 12 | 000000A12345 |

|  |  |  |
| --- | --- | --- |
| Organization sending this information | character | ANAFI |
| UUID 6 | alphanumeric 36 | assigned automatically by gxprep.py program |

|  |  |  |
| --- | --- | --- |
| Genotyping laboratory 7 | character | Weatherbys Ireland |
| Sample ID | alphanumeric | S1234WI2001 |

|  |  |  |
| --- | --- | --- |
| Additional | for future reference |  |
| Array identifier 8 | alphanumeric | 54609\_a |

|  |  |  |
| --- | --- | --- |
| AB – Genotype for SNP Index 1 10 | integer 1 | 0 |
| AB – Genotype for SNP Index 2 10 | integer 1 | 1 |
| AB – Genotype for SNP Index … 10 | integer 1 | 2 |
| AB – Genotype for SNP Index n 9,10 | integer 1 | 5 |

1. Record type is always 706 for this File Format
2. ISO 3166-1 alpha-3 codes (3 characters, capital letters)
3. Breed of evaluation (3 characters, capital letters, BSW, GUE, HOL, JER, RDC, SIM)
4. Breed of animal (3 characters, capital letters)
5. Alpha-numerical, Interbull standard, always 12 characters long
6. UUID, one for every uploaded genotype sequence. Additional information about generation of UUID can be found at <https://wiki.interbull.org/public/uuid?action=print>
7. Genotyping laboratory, among the ones listed in the "Laboratories" table available in the GenoEx-GDE System Data page <https://genoex.org/display>. If the laboratory is not listed in the table, a request should be sent to GenoEx@slu.se
8. Array identifier, one of the listed in the "SNP Arrays" table ("Code" column) available in the GenoEx-GDE System Data page <https://genoex.org/display>.
9. n is equal to the number of SNPs reported in the stem of the Array identifier
10. coded SNP values written as a continuous string.
Acceptable values depend on the Illumina coded allele values, according to the following:

BB→0

AB→1

AA→2

‘unknown’→5

**706 example**

706;BSW;ITA;M;000000A12345;ANARB;09c98b1e-6af8-4254-9768-58d7cd1ddafd;Weatherbys Ireland;S1234WI2001;;54609\_a;021010…

**706 upload**

In order to upload genomic data to GenoEx-GDE database, the file in 706 format has to be zipped. Zip format is the only one accepted by GenoEx-GDE database for the upload. Zip file can be created using the support program gxprep.py as described here: <https://wiki.interbull.org/public/GDE_gxprep_manual#zip_example>. If the User creates the zip file without using the support program gxprep.py, please make sure that the file containing genotypes in 706 format and compressed in the zip archive has name **genotypes.csv**.