ICAR Parentage Analysis Accreditation for DNA Data Interpretation Centres

- Applicant's Guide -

Overview

This document serves as a guide for organizations that submit an application to ICAR for its Parentage Analysis Accreditation for DNA Data Interpretation Centres. In brief, the following are the steps of this ICAR accreditation service:

1. Applicant completes the required forms available on the ICAR web site and submits them to ICAR along with payment of the required fee established by ICAR.

2. ICAR will handle the administration of the application process and once completed, ICAR will inform the Interbull Centre to proceed with the technical component of the accreditation process.

3. Interbull Centre sends by email this Applicant's Guide and the various data files to the contact person provided by the applicant. For Parentage Verification accreditation, these data files include:

   a) Animal genotypes based on the SNP approved by ICAR for parentage verification (≈200)
   b) Pedigree information for animals in the genotype file and potential ancestors
   c) Summary file of Parentage Verification analysis results that must be achieved by the applicant to be successful for ICAR accreditation.

4. Applicant receives the accreditation test files from the Interbull Centre and processes them using internal parentage analysis software. This software must strictly follow the *ICAR Guidelines for Parentage Verification and Parentage Discovery Based on SNP Genotypes,* as posted on the ICAR web site, in order to successfully pass ICAR accreditation.

5. Within 90 days from receiving the accreditation test files from the Interbull Centre, the applicant must send by email to the Interbull Centre, the detailed Results file derived from applying its internal parentage analysis software. Failure to respect the 90-day time period will result in ICAR notification of an unsuccessful accreditation status and the applicant has the option to reapply by initiating step 1 described above at any point in time.

6. Interbull Centre will compare the applicant's Results file to the correct answer file associated with the accreditation test files sent to the applicant. An exact match of results is required to achieve a successful status for receiving the associated ICAR accreditation certificate.

7. Interbull Centre will communicate the Pass or Fail status of the applicant's Results file directly and only to ICAR. As soon as possible thereafter, ICAR administration will issue the proper correspondence to the applicant. In the event of a successful status, ICAR will provide the appropriate Parentage Analysis Accreditation Certificate, which must be renewed, with payment, every two years prior to the anniversary date of the initial date of certification. In the event the applicant was not successful in receiving the ICAR accreditation, the applicant may reapply by initiating step 1 described above at any point in time.

File Names and Description

As outlined above, the process for obtaining ICAR Parentage Analysis Accreditation for DNA Data Interpretation Centres involves the applicant receiving data files from the Interbull Centre and returning a file with results from carrying out the parentage analysis using their internal procedures. The following is a more comprehensive description of the data files involved.

Data Files Received by Applicant

For each applicant, Interbull Centre provides a unique set of data files for testing the parentage analysis procedures developed and used internally by the applicant. Each set of files includes a unique 5-digit numerical value, expressed as XXXXX in the file names presented below. As much as possible the file formats used for this accreditation process follow field descriptions and formats used for other services provided through the Interbull Centre. There are three files involved:

a) Animal Genotypes (genotype_verifyAB_XXXXX.csv)

The .csv file of genotypes follows the same format as one of the genotype exchange file formats to be used by the GenoEx-PSE service (File 704_AB format) as described below. For applicants that prefer to receive the genotype file in the File704_TOP format) to be available for GenoEx-PSE, you are asked to contact the Interbull Centre.

Field Name	Size	Example	Description
Record Type	6 or 7 (fixed)	704-AB (or 704-TOP)	Alphanumeric
Animal ID - Breed Code	3 (fixed)	BSW	Alpha
Animal ID - Nation Code	3 (fixed)	AUS	Alphanumeric (USA, 840, CAN, GBR, FRA, DEU, ...)
Animal ID - Sex Code	1 (fixed)	M	Alpha (M or F)
Animal ID - Registration	18 (maximum)	A1234567890	Alphanumeric
SNP Name	70 (maximum)	ARS-BFGL-NGS-104897	Alphanumeric
Allele 1	1 (fixed)	A/B or A/C/G/T	Alpha
Allele 2	1 (fixed)	A/B or A/C/G/T	Alpha

b) Pedigree (ped_verify_XXXXX.csv)

The file of pedigree data follows the same format as used by Interbull for distributing pedigree information to its customers (pedig_BBB.csv, where BBB refers to the Interbull 3-character breed code). For this ICAR accreditation process, the following layout and description is used.

Field Name	Size	Example	Description
Animal ID (note 1)	19 (fixed)	XXXZZZF001234567890	Alphanumeric
Sire ID (note 1)	19 (fixed)	XXXZZZM002345678901	Alphanumeric (filled as 0000000000000000000 when unknown)
Dam ID (note 1)	19 (fixed)	XXXZZZF003456789012	Alphanumeric (filled as 0000000000000000000 when unknown)
Animal's Birth Date	8 (fixed)	YYYYMMDD	Numeric
Animal Name	Variable		Always blank
Verification Status	1 (fixed)	U	Always filled as U
Sending Country	3 (fixed)	ZZZ	Always filled as ZZZ

Note 1: First six characters for "Breed" and "Country of Registration" are filled with fictitious values of XXX and ZZZ respectively.

In essence, the applicant may only use the first four fields included in this comma delimited .csv file and ignore the last three fields. Note that within the Animal ID field, the sex of the animal (i.e.: M or F as character 7), as well as the Animal's Birth Date, are important content that is required for conducting correct parentage analysis, especially for Parentage Discovery.

The Pedigree file includes a single record for each of the approximately 15,000 animals, of which approximately 4,000 represent the animals/progeny of interest and the remainder are potential parents of those animals. The progeny/animals of importance for this ICAR Accreditation are those that have an Animal ID beginning with either XXXZZZF00142 for females or XXXZZZM00142 for males, and these animals must all be included in the Results File to be returned by the applicant to the Interbull Centre.

For Parentage Verification accreditation, the Animal Genotype file will have up to 200 records (i.e.: one for each called SNP) for each animal in the Pedigree file, which means a maximum total of 3M records.

c) Summary File (summary_XXXXX.csv)

The ICAR accreditation for Parentage Verification tests 22 distinct cases, as outlined in the following table, depending on the result found for verifying the recorded sire, dam and mating combination of sire by dam. Specific to the genotype and pedigree files provided to the applicant the counts of N1, N2, ..., N22 are provided in the Summary File to enable the applicant to know for certain when they have not yet passed the accreditation process. Note that the 22 cases tested by this ICAR accreditation process represent a selected subset of all possible cases that should be handled internally by the applicant for parentage analysis. When the internal parentage analysis procedures applied by the applicant have correctly applied the ICAR Guidelines, then the Results File (see below) should provide counts that match exactly those in the Summary File.

Case No.	Sire Conclusion	Dam Conclusion	Mating Conclusion	Count
1	Sire_Accepted	Dam_Accepted	Mating_Accepted	N1
2	Sire_Accepted	Dam_Accepted	Mating_Doubtful	N2
3	Sire_Accepted	Dam_Accepted	Mating_Excluded	N3
4	Sire_Doubtful	Dam_Accepted	Not_Checked	N4
5	Sire_Excluded	Dam_Accepted	Not_Checked	N5
6	Sire_Accepted	Dam_Doubtful	Not_Checked	N6
7	Sire_Accepted	Dam_Excluded	Not_Checked	N7
8	Sire_Doubtful	Dam_Doubtful	Not_Checked	N8
9	Sire_Doubtful	Dam_Excluded	Not_Checked	N9
10	Sire_Excluded	Dam_Doubtful	Not_Checked	N10
11	Sire_Excluded	Dam_Excluded	Not_Checked	N11
12	Sire_Unknown	Dam_Accepted	Not_Checked	N12
13	Sire_Accepted	Dam_Unknown	Not_Checked	N13
14	Low_Call	Low_Call	Not_Checked	N14
15	Low_Call	Dam_Accepted	Not_Checked	N15
16	Low_Call	Dam_Doubtful	Not_Checked	N16
17	Low_Call	Dam_Excluded	Not_Checked	N17
18	Sire_Accepted	Low_Call	Not_Checked	N18
19	Sire_Doubtful	Low_Call	Not_Checked	N19
20	Sire_Excluded	Low_Call	Not_Checked	N20
21	Sire_Not_Genotyped	Dam_Accepted	Not_Checked	N21
22	Sire_Accepted	Dam_Not_Genotyped	Not_Checked	N22

Note that Case 14 represents the situation when the genotype for the animal itself has too low of a call rate to be used for parentage analysis, so the conclusion of Low_Call is assigned for the Sire, Dam and the Mating.

Results File to be Returned to the Interbull Centre

Once the applicant obtains output from their internal parentage analysis procedures for which the counts exactly match those for each case included in the Summary File, then the applicant should create and send a detailed Results File back to the Interbull Centre. **This file must be a comma (,) delimited .csv file respecting the naming structure as follows**:

results_XXXXX.csv

where XXXXX refers to the same 5-digit numerical value included in the data file provided to the applicant by the Interbull Centre.

The content of the Results File should respect the file layout described below and only include those animals (≈ 4,000) in the genotype file that have the numerical portion of their Animal ID (Animal ID - Unique Number) starting with "00142".

Field No.	Field Name	Size	Example	Description
1	Animal ID - Breed Code	3 (fixed)	XXX	Always XXX
2	Animal ID - Country of Registration	3 (fixed)	ZZZ	Always ZZZ
3	Animal ID - Sex Code	1 (fixed)	F	Always F or M
4	Animal ID - Unique Number	12 (fixed)	001426789012	Numeric and must start with 00142
5	No. Usable SNP - Animal	3 (maximum)	190	Count of usable SNP in the Animal's genotype
6	No. Usable SNP - Sire	3 (maximum)	196	Count of usable SNP in the Sire's genotype
7	No. Usable SNP - Dam	3 (maximum)	192	Count of usable SNP in the Dam's genotype
8	Sire Conclusion	Variable	Sire_Accepted	Must be assigned as listed in the Summary File table above under "Sire Conclusion"
9	Dam Conclusion	Variable	Dam_Excluded	Must be assigned as listed in the Summary File table above under "Dam Conclusion"
10	Mating Conclusion	Variable	Mating_Excluded	Must be assigned as listed in the Summary File table above under "Mating Conclusion"
11	No. SNP Conflicts - Sire[1]	3 (maximum)	5	Count of the number of SNP in conflict between the Animal and the recorded Sire
12	No. SNP Conflicts - Dam[2]	3 (maximum)	4	Count of the number of SNP in conflict between the Animal and the recorded Dam
13	No. SNP Conflicts - Mating[3]	3 (maximum)	5	Count of the number of SNP in conflict when considering the mating of the recorded Sire and Dam as the parents of the Animal
14	Case Number	2 (maximum)	8	Fill in with the appropriate Case No. as described in Summary File section above

Note 1: Field 11: No. SNP Conflicts - Sire must be blank when Sire Conclusion has a result of Low_Call.
Note 2: Field 12: No. SNP Conflicts - Dam must be blank when Dam Conclusion has a result of Low_Call.
Note 3: Field 13: No. SNP Conflicts - Mating must be blank when Mating Conclusion has a result of Not_Checked.

For sake of clarity, the following are examples of correct records with required fields having a value and the three fields associated with No. SNP Conflicts having missing values when appropriate:

XXX,ZZZ,M,00142xxxxxxx,187,192,0,Sire_Accepted,Dam_Not_Genotyped,Not_Checked,0,,,22
XXX,ZZZ,F,00142xxxxxxx,196,196,192,Sire_Doubtful,Dam_Accepted,Not_Checked,5,0,,4
XXX,ZZZ,M,00142xxxxxxx,195,196,196,Sire_Accepted,Dam_Accepted,Mating_Doubtful,0,0,4,2
XXX,ZZZ,F,00142xxxxxxx,179,196,196,Low_Call,Low_Call,Not_Checked,,,,14

Assessment of Accreditation Results

Once the Results File is sent by the applicant to the Interbull Centre, it will be compared, record by record, to the correct answer key associated with the specific set of accreditation test files originally sent to the applicant. Only in the event that this comparison yields identical results will ICAR subsequently grant a successful accreditation status and certification. If this is not the case, further options available to the applicant are described in the Overview section of this document. If desired, the applicant may request to receive from the Interbull Centre the correct answer key for the Results File, which would help identify the specific areas that need improvement for achieving a successful outcome if the applicant decided to re-apply for the ICAR Parentage Analysis Accreditation for DNA Data Interpretation Centres.