



THE GLOBAL STANDARD
FOR LIVESTOCK DATA

ICAR Guidelines for Parentage Verification and Parentage Discovery Based on SNP Genotypes

Prepared by: ICAR DNA Working Group
Revised May 2023

Purpose

In 2018 ICAR introduced two new services related to the use of SNP genotypes for dairy and beef cattle. One of these is the **Genotype Exchange Parentage SNP Exchange (GenoEx-PSE)** Service, which allows countries to exchange SNP genotypes for the purpose of offering parentage analysis services to specific breed populations at the national level. Another new service, which is also a prerequisite for GenoEx-PSE, is the service of ICAR SNP-Based Parentage Analysis Accreditation for Data Interpretation Centres, which allows organizations wishing to carry out SNP-based parentage analysis services to be accredited by an independent third party. Organizations wishing to be a service user of GenoEx-PSE must first receive ICAR Accreditation for Data Interpretation Centres.

In July 2012, the International Society for Animal Genetics (ISAG) approved standards for genotyping laboratories to conduct parentage verification. These *"Guidelines for cattle parentage verification based on SNP markers"* were based on carrying out parentage verification using a set of 200 recommended SNP, with 100 considered as core SNP and a second set of 100 as backup SNP.

The purpose of this document is two-fold:

1. Revise the original ISAG guidelines for conducting parentage verification in dairy and beef cattle using the 200 SNP recommended by ISAG, and
2. Establish new guidelines for conducting parentage discovery based on the SNP included for this purpose in the GenoEx-PSE genotype exchange service, which initially totals 554.

Principles

When carrying out parentage analysis, which includes parentage verification and/or parentage discovery, the following underlying principles should be considered:

- A consistent set of SNP must be defined for use by all organizations for international recognition of parentage analysis accreditation and for the subsequent delivery of "certified" parentage information.
- Each animal involved in the parentage analysis process (i.e.: animal and each potential parent) must have a SNP genotype available for which a minimum proportion of the defined set of SNP have been called and are available. When establishing such minimum requirements consideration should be given to the inclusion of the defined SNP on various genotyping chips used widely in the population of animals being considered.
- While it is understood that only informative SNP (i.e.: SNP whereby the animal and the parent in question are both homozygous) provide useful information for parentage analysis, it is more practical that guidelines are based on the total number of SNP available for the animal and parent(s) in question. Roughly speaking, about one-third of available SNP for parentage analysis are informative but this proportion depends on the average minor allele frequency of the included SNP within the population of animals being considered.

Parentage Verification

Since ICAR introduced the service of SNP-Based Parentage Analysis Accreditation for Data Interpretation Centres in 2018, the following guidelines have been used as the international standard for ICAR accreditation:

1. Upon approval by the ICAR DNA Working Group, specific individual SNP included in the current group of 200 SNP recommended by ISAG for parentage verification in cattle, may be deemed inappropriate for inclusion. This reduction in the total set of SNP to be used would be applied by all organizations receiving ICAR Accreditation for Data Interpretation Centres. Annex 1 attached lists the SNP that have shown to cause problems for parent verification and therefore are recommended for exclusion from the original list of 200 recommended by ISAG for parentage verification. On an ongoing basis ISAG may identify other problematic SNP from various chip platforms and technologies for inclusion in Annex 1 upon final approval by the ICAR DNA Working Group.
2. Given the improved accuracy of parentage verification achieved by the inclusion of more SNP, the original two-step process suggested by ISAG should be replaced by a single analysis based on the full set of approximately 200 SNP for parentage verification, excluding those in Annex 1.
3. The required minimums in terms of number of SNP, as outlined in the current ISAG guidelines, must be scaled to reflect the total number of SNP to be used for parentage verification analysis. For example, if the total number of SNP from those recommended by ISAG, is reduced from 200 to 195 (as is the current situation), then the minimum number of SNP available in the profile of each animal and potential parent must be scaled to 185 (i.e.: 95% truncated down), based on the ISAG policy of applying a minimum of 95 out of 100.
4. For assigning the parentage verification status according to the number of SNP conflicts found, the revised rules approved by ICAR are the following:

Step 1: Conduct a separate verification for each combination of the animal with its recorded sire and/or dam with a SNP genotype. The informative SNP are those for which the animal and reported parent are both homozygous and a conflict is considered when they are each homozygous for a different allele for any informative SNP. Based on the minimum criteria of 185 SNP available for the animal and each parent, the minimum number of common SNP available for verifying each animal-parent combination is 175 (i.e.: $195 - (2 \times (195-185)) = 175$).

For this step, the following rules apply for assigning the parentage verification status:

- Number of mismatches/SNP conflicts: 0 - 2 => Parent Accepted
- Number of mismatches/SNP conflicts: 3 - 5 => Parent Doubtful
- Number of mismatches/SNP conflicts: >5 => Parent Excluded

Step 2: In the case that both sire and dam have a status of "Parent Accepted" from Step 1, verify that the combination of those parents is acceptable. In this case the informative SNP are those for which both verified parents are homozygous and the progeny is heterozygous. A conflict exists when the parents are homozygous for the same allele at any informative SNP while the progeny is heterozygous. In this case, the minimum number of common SNP available is 165 ($195 - (3 \times (195-185)) = 165$).

For this step, the following rules apply for confirming the parentage verification status for the combination of verified parents:

- Number of mismatches/SNP conflicts: 0 - 3 => Mating Accepted
- Number of mismatches/SNP conflicts: 4 - 7 => Mating Doubtful
- Number of mismatches/SNP conflicts: >7 => Mating Excluded

5. For animals with only one parent genotyped, only those animal-parent combinations achieving the status of "Parent Accepted" from Step 1 would qualify for the organization to issue an official confirmation of parentage for that parent. For animals with both parents genotyped, only those animals achieving the status of "Parent Accepted" from Step 2 would qualify for the organization to issue an official confirmation of parentage including both parents.
6. As an added service for those organizations receiving ICAR accreditation to carry out parentage discovery, the process outlined below could be applied to all animals for which the parentage verification result was either "Parent Doubtful" or "Parent Excluded" in either Step 1 or Step 2 above. Any animal that received the parentage verification result of "Parent Accepted" for both parents as well as "Mating Accepted" for the combination of parents must be excluded from any subsequent parentage discovery analysis.

Parentage Discovery

In 2023, ICAR will expand its current services of SNP-Based Parentage Analysis Accreditation for Data Interpretation Centres to include parentage discovery as a separate option in addition to parentage verification. Prior to this service, no international guidelines existed for organizations to carry out parentage discovery even though most, if not all, genetic evaluation service providers have developed such processes internally. As with parentage verification, the accuracy of parentage discovery is improved as the number of SNP included increases. For the GenoEx-PSE service, a list totalling 554 SNP has been defined for the genotype exchanges involving service users that have been accredited by ICAR for this level of parentage analysis and have agreed to upload these SNP to the GenoEx-PSE database at the Interbull Centre, which is the requirement for downloading the same. These 554 SNP include the 200 SNP recommended by ISAG for parentage verification in cattle as well as an additional group of 354 SNP. In addition to the 200 SNP for parentage verification another 75 for parentage discovery are spread across chromosomes 1 to 29 while the remaining 279 SNP were selected from only ten chromosomes, specifically 1, 2, 3, 5, 7, 8, 11, 13, 19 and 21. This strategy for SNP selection was adopted to reduce the accuracy of genotype imputation and genomic predictions in the event that any GenoEx-PSE service user attempts to use the exchanged genotype in this manner even though it is clearly prohibited as outlined in the GenoEx-PSE Service Agreement.

To be consistent with the principles and revised guidelines for parentage verification outlined above, the following are guidelines approved by ICAR for parentage discovery:

1. The ICAR DNA Working Group may, from time to time, identify and approve SNP, from among those included in the GenoEx-PSE service, that must be excluded for carrying out parentage discovery, which are listed in Annex 2 attached. Any such SNP would include those approved for exclusion for parentage verification and may also include other SNP once there is sufficient reason to do so.
2. Organizations carrying out parentage discovery services must implement quality assurance procedures that ensure the following:
 - That a discovered parent is older than the animal and, in fact, not an offspring. Cattle normally reach sexual maturity at 11-12 months of age, but this can be as young as 8-9 months, and even younger if in-vitro fertilization is a technology used within the population. Under normal circumstances, a minimum of 17 months between the dates of the animal and its predicted parent is recommended to ensure that the predicted parent could have been sexually mature at the time of the breeding. For the ICAR parentage analysis accreditation services, a minimum age difference of 517 days is applied for allow for qualifying parents.
 - That a discovered parent is of the appropriate sex such that sires are male and dams are female.
 - That procedures must include methods for identifying and listing genetically identical animals. While this can be done as a separate processing step, the ICAR accreditation for parentage discovery will test that such animals are identified within the processing procedures such that a discovered parent is also reported as having known genetically identical siblings. For ICAR parentage analysis accreditation services, any pair or group of animals should be considered as genetically identical

when their genotypes have $\leq 2\%$ differences among the common set of SNP that were successfully called in their genotypes. For parentage verification, which is currently based on 195 SNP, this means no more than 3 called SNP in common that are not the same and for parentage discovery, which is currently based on 549 SNP, this means no more than 10 called SNP in common that are not the same.

3. Based on an assessment of SNP lists associated with various SNP chips used internationally to genotype dairy and/or beef cattle, essentially all chips have at least 450 SNP in common with the 554 SNP recommended for parentage discovery. The only two exceptions to-date are the Illumina Bovine3K BeadChip with 233 SNP in common, which is no longer in use and should not be included for any parentage discovery analysis, and the Affymetrix Swiss LD Version 1, with 346 SNP in common. Given possible call rates of genotypes for the animal and any potential parent to be discovered, it has been approved by ICAR that each genotype included in such an analysis has a minimum of 400 of the 554 SNP in common between the animal and its candidate parent in order to conduct parentage discovery as well as a minimum of 400 SNP in common between the trio of the animal with the candidate parents after discovery when confirming the mating combination.
4. Given that genotyping SNP chips actively being used in cattle populations globally have a varying number of the 554 SNP defined for inclusion in the GenoEx-PSE service, parentage discovery results must be based on a percentage of SNP available between the animal and any potential parent being considered. The following is recommended for assigning the parent discovery status:

Step 1: In separate processes, attempt to discover either the **sire** (i.e.: the male older than the animal with the fewest conflicts) or **dam** (i.e.: the female older than the animal with the fewest conflicts) of the animal based on SNP genotypes available. Based on the minimum criteria for each SNP genotype to be included, as outlined in point 3 above, the minimum number of common SNP between the animal and each candidate parent must be 400.

The informative SNP used for the processing to discover either the sire or dam are those for which the animal and a candidate parent are both homozygous (i.e.: AA or BB) and a conflict is considered when they are each homozygous for a different allele for any informative SNP (i.e.: one is AA and the other is BB).

For this step, the following rules apply for assigning the status of each parent discovered:

- Percentage of 400+ common SNP with a conflict: 0 to $\leq 0.5\%$ => Parent Discovered
- Percentage of 400+ common SNP with a conflict: >0.5 to $\leq 2.0\%$ => Parent Possible
- Percentage of 400+ common SNP with a conflict: $>2.0\%$ => Parent Excluded

Consider the following example for the calculation of percentage SNP with a conflict:

SNP Number	1	2	3	4	5	6	7
Candidate 1:	AA	AA	AB	BB	--	BB	AB
Candidate 2:	--	BB	AB	BB	AA	AB	AA

This example has 5 informative SNP for parentage analysis, which are numbers 2, 3, 4, 6 and 7. SNP numbers 1 and 5 are excluded since these SNP were not successfully called for one of the animals involved. SNP number 2 shows a conflict since the two candidates are both homozygous but for a different allele. Therefore, the percentage of conflicts is calculated as $1/5 \times 100 = 20\%$.

For animals for which the discovery process led to all qualifying results as "Parent Excluded", the discovery conclusion for that animal is considered as Parent Not Found.

For animals that do not have enough informative SNP available to be included in the analysis and/or are found during the discovery process that there is no candidate parent with at least 400 SNP in

common, the discovery conclusion outcome should be Parent Not Found instead of any of Discovered, Possible or Excluded as outlined above.

Step 2: In the case that an animal has both a sire and dam with a successful status of "Parent Discovered" from Step 1 or one parent qualified as "Parent Discovered" and the other parent had the "Accepted" result during parentage verification, this parent combination must also be checked. The informative SNP used to confirm such trios of an animal and two parents are those for which both parents that were verified or discovered are homozygous (i.e.: AA or BB) and the progeny is heterozygous (i.e.: AB). A conflict exists when the parents are homozygous for the same allele at any informative SNP (i.e.: both are AA or both are BB) while the progeny is heterozygous (i.e.: AB).

For this step, the following rules apply for assigning the status of the combination of parents discovered, which requires the trio of the animal with both of the candidate parents after discovery to have at least 400 SNP in common:

- | | | |
|--|-------------------------|---------------------|
| - Percentage of 400+ common SNP with a conflict: | 0 to $\leq 1.0\%$ | => Mating Confirmed |
| - Percentage of 400+ common SNP with a conflict: | > 1.0 to $\leq 4.0\%$ | => Mating Possible |
| - Percentage of 400+ common SNP with a conflict: | $> 4.0\%$ | => Mating Excluded |

For trio combinations of animal with its candidate parents that do not have at least 400 SNP in common, the discovery conclusion outcome for the mating should be Mating Not Checked instead of any of Confirmed, Possible or Excluded as outlined above.

5. For animals with only one parent with the status of "Parent Discovered" in Step 1, only that animal-parent combination would qualify for the organization to issue an official confirmation of parentage for that parent. For animals with both parents with the status of "Parent Discovered" in Step 1, only those animals achieving the status of "Mating Confirmed" from Step 2 would qualify for the organization to issue an official confirmation of parentage with the status of "Parent Accepted" for both parents.

ANNEX 1: SNP Among the ISAG 200 to be Excluded from Parentage Verification for ICAR Accreditation

SNP Name (Illumina Bead Chips)	ISAG Group	Reason for Exclusion
ARS-USMARC-Parent-DQ837645-rs29015870	Core	Clustering issues*
ARS-USMARC-Parent-DQ786766-rs29012070	Core	Clustering issues*
ARS-BFGL-NGS-76191	Backup	Clustering issues*
BTA-100621-no-rs	Backup	Clustering issues*
ARS-BFGL-NGS-99210	Backup	Tri-allelic**

* - McClure et al. (2015)

** - Based on sequence validation to be specifically problematic with bead chips

SNP associated with other chip platforms and/or technologies may be added over time.

ANNEX 2: SNP Included in GenoEx-PSE to be Excluded from Parentage Discovery for ICAR Accreditation

SNP Name (Illumina Bead Chips)	ISAG Group	Reason for Exclusion
ARS-USMARC-Parent-DQ837645-rs29015870	Core	Clustering issues*
ARS-USMARC-Parent-DQ786766-rs29012070	Core	Clustering issues*
ARS-BFGL-NGS-76191	Backup	Clustering issues*
BTA-100621-no-rs	Backup	Clustering issues*
ARS-BFGL-NGS-99210	Backup	Tri-allelic**

* - McClure et al. (2015)

** - Based on sequence validation to be specifically problematic with bead chips

SNP associated with other chip platforms and/or technologies may be added over time.