CDCB Policy

QUICK DISCOVERY OF CLOSE RELATIVES (QDisc) POLICY

This policy for creating a quick close relative discovery service based on ICAR-approved 550 SNP subset was approved by the CDCB Board of Directors on 12/14/2022. Last updated in 03/07/2023

0. REVISIONS

March 2023: Inclusion of wording relative to the End User Agreement and the use of results.

1. POLICY STATEMENT

The Council on Dairy Cattle Breeding (CDCB) acknowledges a growing demand for genomic services on commercial herds with limited information on animal pedigrees and with the main goal of using genomic predictions as a selection and culling tool for young female calves. Since 2019 CDCB offers a quick turnaround genomic prediction service using approximated methods with lower accuracy than the existing CDCB genomic services, providing a limited service and only including results on a limited number of traits. However, commercial herds are often also interested in improving their pedigree information using genomics. Currently, quick turnaround users are required to nominate their animals and load the genotypes into the database, which creates unnecessary burden to the user and usage of the database. This policy describes a service that allows all nominators (irrespectively of their participation to the quick turnaround evaluation service) to receive a quick discovery of close relatives based on a reduced SNP set.

2. POLICY DETAILS

2.1 REQUIREMENTS AND CHARACTERISTICS
- All CDCB certified genomic nominators can access the service. **The service is set up upon request**.
- The service is based on the recommended 550 ICAR SNP subset to be used for parentage discovery.
- Genotypes submitted by the requesting nominator will need to follow the [CDCB accepted genotype file format](#).
  - Only FinalReport files are *required*. SampleSheet files *can* be used to submit a differential animal ID (in the Sample_Name field).
  - Although any number of SNPs can be submitted in the file, only the aforementioned 550 ICAR SNP subset will be retained in the analysis.
  - The service is chip-agnostic.
  - The FinalReport file must contain called SNP for at least 300 of the 550 SNP being used.
- Incoming genotype files will be stored for not more than 2 months (for internal investigation reasons) and then automatically removed.

- Unless the user re-submits the files for full service, the genotypes will not be included in the CDCB Cooperator database, shared with third parties, or used for any other purpose than delivering this service.
- The service relies on a specific database (QDiscDB) including the 550 ICAR SNP subset for all the CDCB Collaborator database. Comparisons are made with QDiscDB and with all samples submitted in the file.
- QDiscDB will be updated on a daily basis, and refreshed completely 3 times a year.
- Since submitted samples are not stored in any database, no comparisons across previously submitted files to this service are possible.
- This service relies on the CDCB cooperator database availability (for discovered ID information), so it will not be provided during database maintenance events.

2.2 DELIVERABLES

Results of the quick discovery process will be distributed using simplified formats containing:

- A status file (qdstatus.csv) including: the ID used (Input_ID), the number of usable SNPs (Input_SNP_calls), the number of discovered close relatives (Disc_matches), a Message (optional)
- A results file (qdisc.csv) including: the ID used (Input_ID), the discovered animal (Disc_ID), the discovered animal date of birth (Disc_DOB), the discovered animal sex (Disc_SEX), total SNPs compared (SNP_compared), number of conflicting SNPs (SNP_conflict), a Message (optional).
  - All clones found are reported to facilitate confirmation or misidentification of samples.
  - Only usable genotypes in the CDCB Collaborator Database (use_ind=Y) are reported.
  - A maximum of 5 animals by sex sorted by animal date of birth, will be reported.
- The quick discovery service is delivered once and are never updated. This implies that re-submission of files and/or samples will be considered as a new request and charged accordingly.
- Results are distributed exclusively to the requester in a .zip file.
- Results may be used for internal noncommercial research and internal business purposes. Results containing owner identifying information may be disclosed, represented, or otherwise communicated to the owner of the animals, as long as the substance of such Reports is not materially modified and the user complies with this the End User Agreement (see 2.4) and this policy.

2.3 TURNAROUND

The quick discovery service will be automated and results delivered within 48 hours after the genotype files and, if applicable, accessory files are successfully uploaded into the quick discovery dedicated SFTP area.

2.4 END USER AGREEMENT

Customized reports will be defined in an End User Agreement (EUA) between the CDCB and all collaborators requesting the quick discovery of close relatives service. Unless otherwise defined in the EUA, you may not use or incorporate any information provided by CDCB as part of a competing genetic or genomic prediction or estimate of genetic or genomic merit or traits.

2.5 FEE STRUCTURE

	Quick Discovery fee
Quick discovery of close relatives	$ 1

This service has no setup fee. Samples submitted will be charged the quick discovery fee, irrespectively of the successful discovery of close relatives. The submission of a file implies the acceptance of all the rules here described. **This fee is non-refundable and no credits will be issued, even in case of samples successively sent to CDCB's full-service.**