CDCB's Genotyping Laboratory Certification Program

José A. Carrillo, George Wiggans, Ezequiel Luis Nicolazzi, Kaori Tokuhisa, Duane Norman and João Dürr

Interbull Annual Meeting

June 23, 2019 – Cincinnati, OH

CDCB

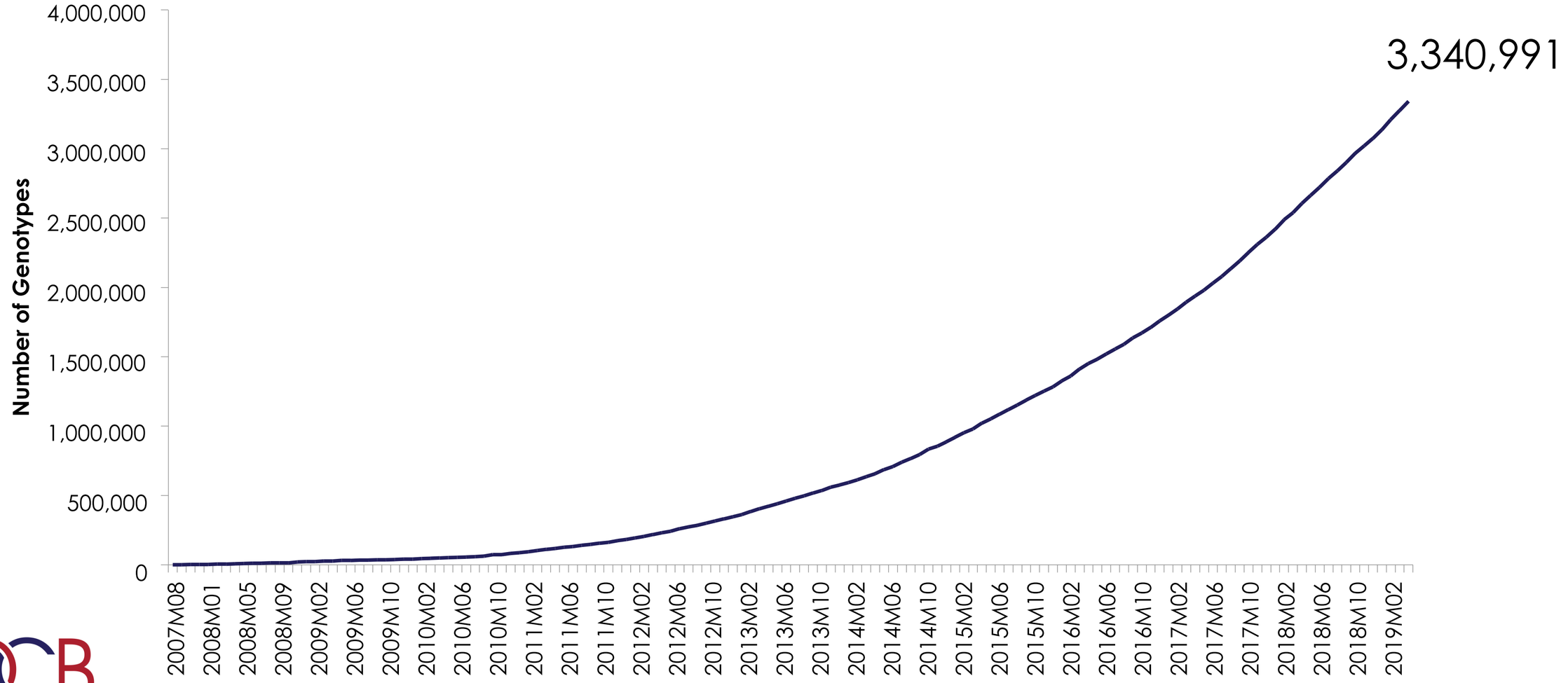
COUNCIL ON DAIRY CATTLE BREEDING

Overview

- Impact of recent technological advances
  - Large number of animals genotyped at a relatively low cost
  - Massive amount of data that should be collected and analyzed
- Results depend on the quality of the input data
  - Data obtained from different types of chips and under two distinct technologies
  - CDCB exchanges data with many different organizations
  - The data must be scaled and standardized to make them comparable
- Maintaining the integrity of the CDCB database requires a robust quality control (QC) program
  - CDCB developed a customized QC program according to its requirements

CDCB
COUNCIL ON DAIRY CATTLE BREEDING

Genotypes in CDCB Database



CDCB
COUNCIL ON DAIRY CATTLE BREEDING

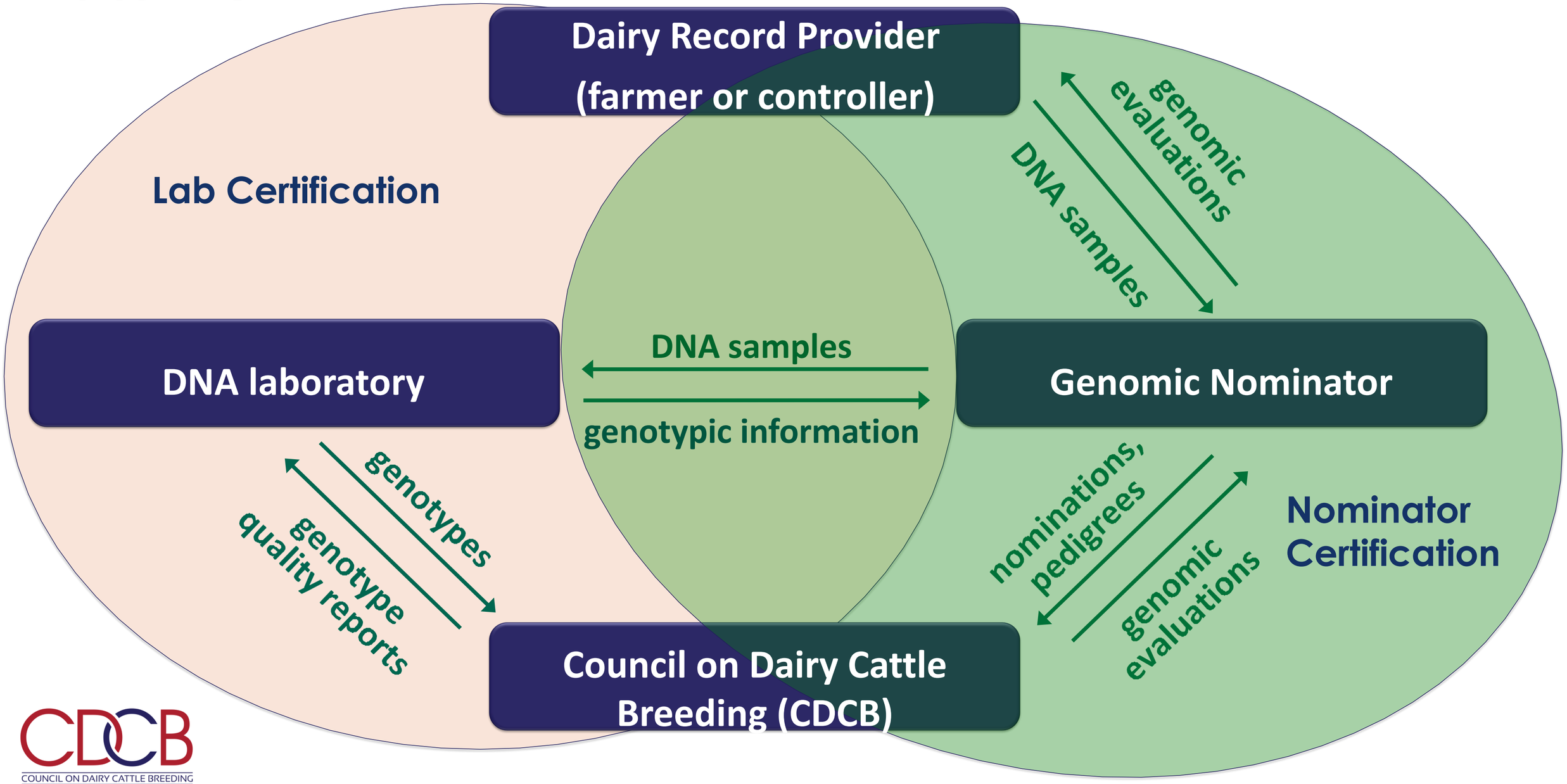
Purpose of the laboratories QC program

Ensure the accuracy and uniformity of all records included in the national genomic evaluation

- Monitor certified laboratories performance regularly to ensure quality of data
- Detect the needs or issues experienced by laboratories
- Advise or find solutions for issues/concerns faced by labs
- Facilitate the exchange of data (in the most efficient way)
- Improve communication between labs and the CDCB

CDCB
COUNCIL ON DAIRY CATTLE BREEDING

Data flow

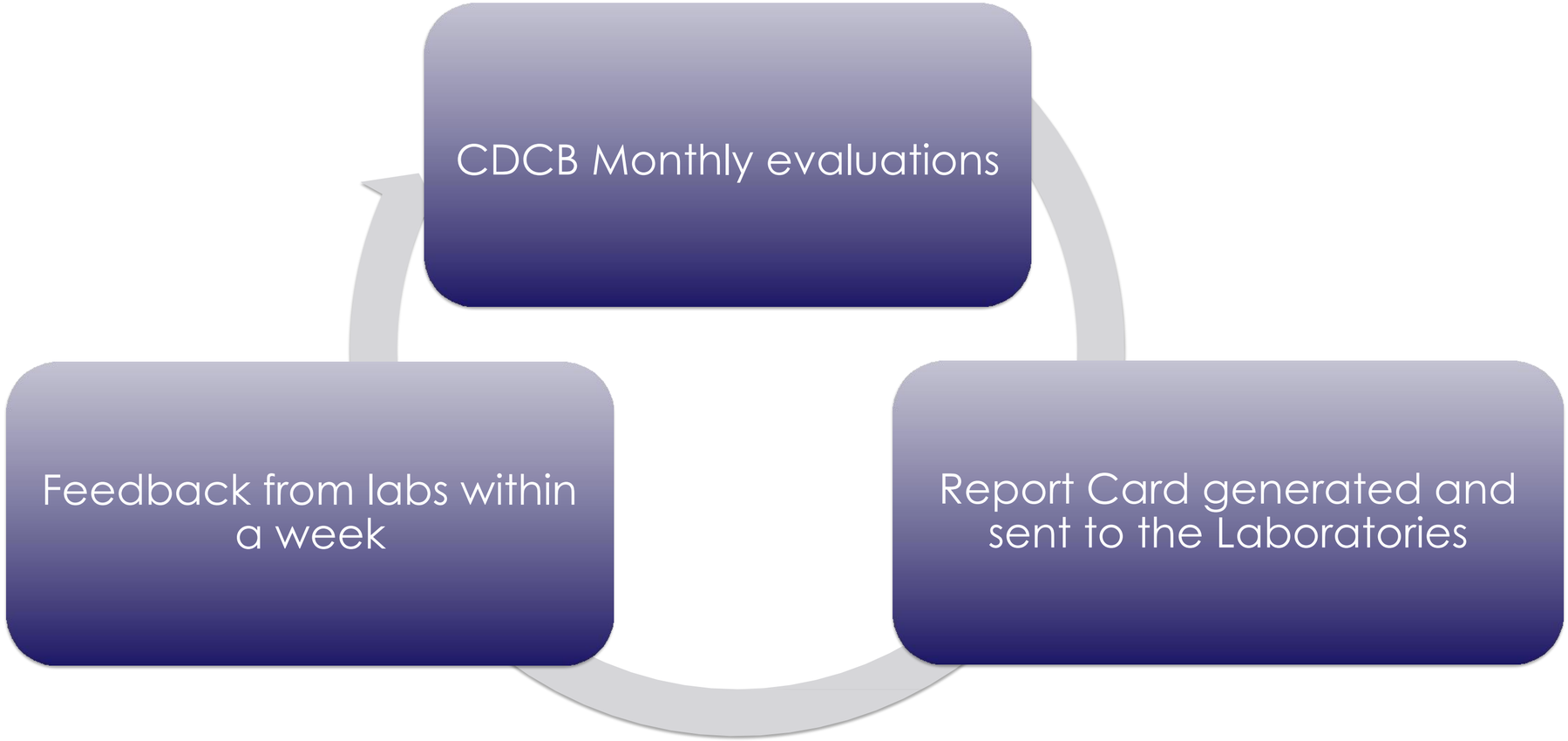


CDCB
COUNCIL ON DAIRY CATTLE BREEDING

CDCB - Certification Process

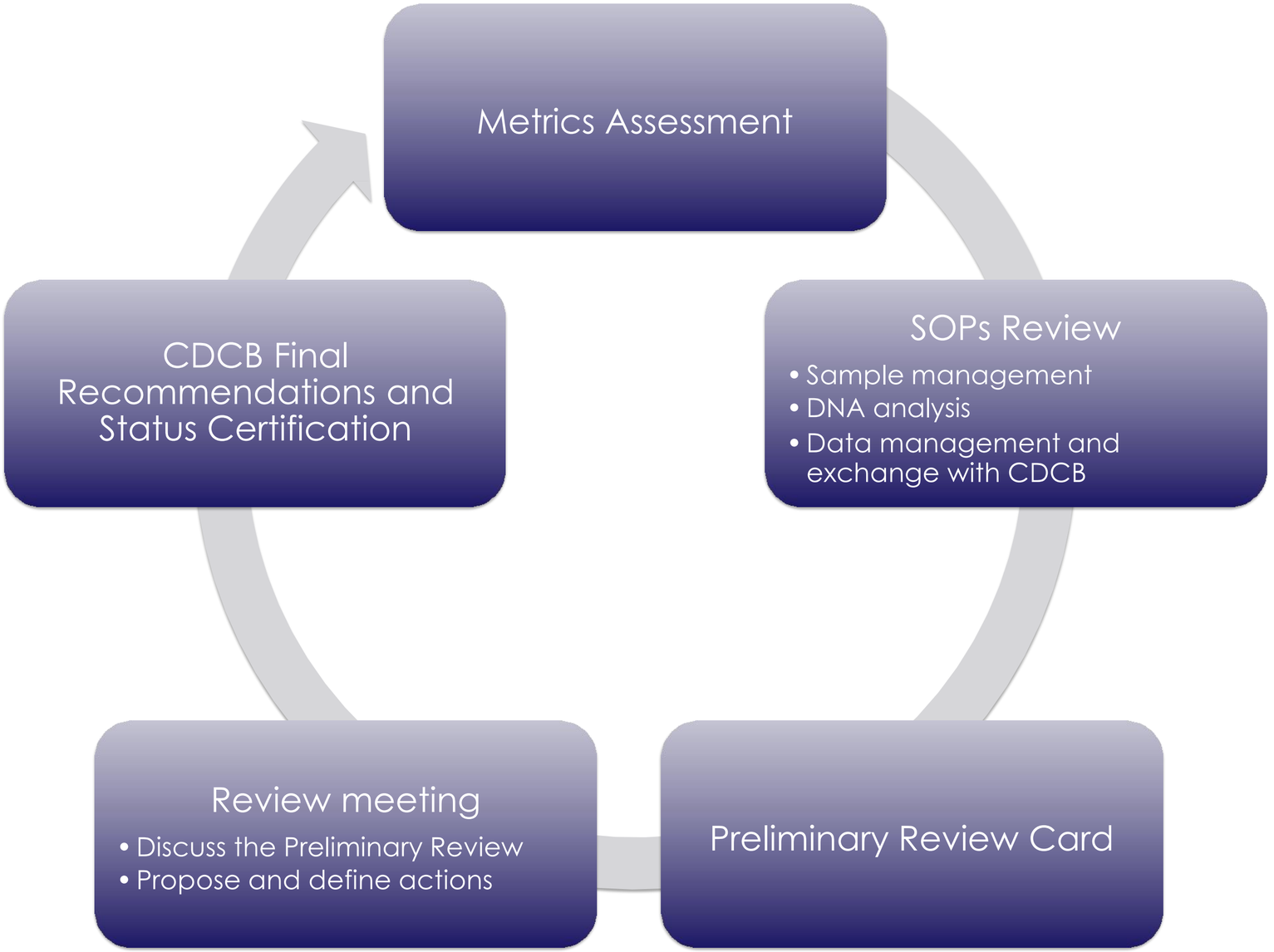
CDCB
COUNCIL ON DAIRY CATTLE BREEDING

Monthly Report Cards



CDCB
COUNCIL ON DAIRY CATTLE BREEDING

Annual Review Plan



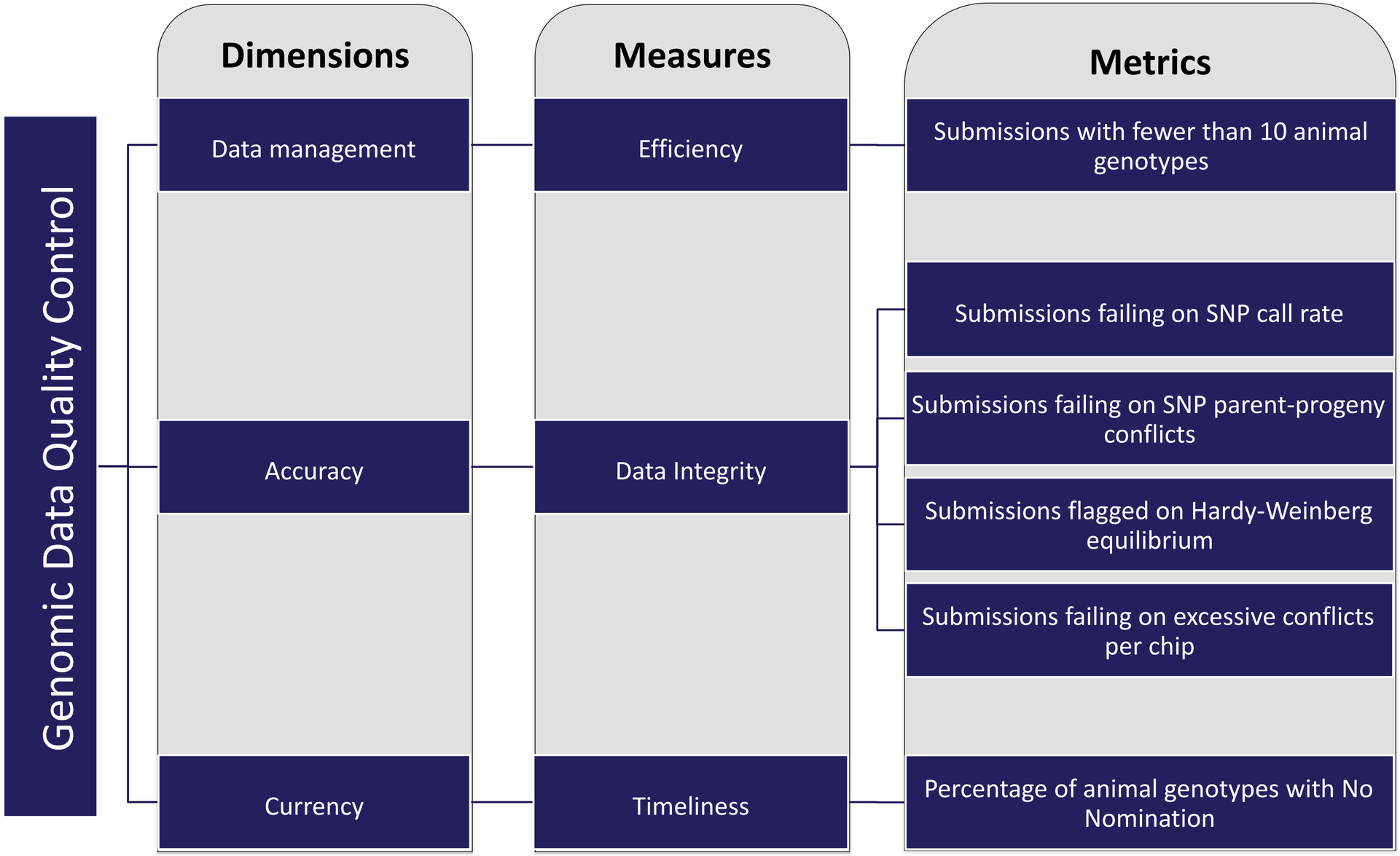
CDCB
COUNCIL ON DAIRY CATTLE BREEDING

QC Metrics in Monthly Report Cards

- Metrics are classified according to their impact on the quality and processing of the data

Critical	Submissions with fewer than 10 animal genotypes (10%)
	Submissions failing on SNP call rate (50%)
	Submissions failing on SNP parent-progeny conflicts (25%)
	Submissions flagged on Hardy-Weinberg equilibrium (HWE) (50%)
Major	Percentage of animal genotypes with No Nomination (3%)
	Submissions failing on excessive conflicts per chip (10%)

CDCB
COUNCIL ON DAIRY CATTLE BREEDING



CDCB
COUNCIL ON DAIRY CATTLE BREEDING

Performance Metrics Assessment

Lab performance metrics for XX

	Jan-18	Feb-18	Mar-18	Apr-18	May-18	Jun-18	Jul-18	Aug-18	Sep-18	Oct-18	Nov-18	Dec-18
Total Files Processed	78	54	50	58	52	59	57	44	60	46	49	58
Submissions with fewer than 10 animal genotypes	15	8	5	4	4	11	2	8	3	5	8	6
Submissions failing on SNP call rate	39	24	25	25	19	30	0	29	42	32	30	37
Submissions failing on SNP parent-progeny conflicts	11	10	9	3	1	2	0	2	0	1	1	3
Submissions failing on HWE	3	4	6	2	4	1	0	5	1	0	3	2
Percentage of animal genotypes with No Nomination	0	0	0	0	0	0	0	0	0	0	0	0
Submissions failing on excessive conflicts per chip	45	27	11	18	10	7	0	13	17	10	13	15

	Thresh	Jan-18	Feb-18	Mar-18	Apr-18	May-18	Jun-18	Jul-18	Aug-18	Sep-18	Oct-18	Nov-18	Dec-18
Submissions with fewer than 10 animal genotypes	10%	19.23%	14.81%	10.00%	6.90%	7.69%	18.64%	4.23%	18.18%	5.00%	10.87%	16.33%	10.34%
Submissions failing on SNP call rate	50%	50.00%	44.44%	50.00%	43.10%	36.54%	50.85%	61.97%	65.91%	70.00%	69.57%	61.22%	63.79%
Submissions failing on SNP parent-progeny conflicts	25%	14.10%	18.52%	18.00%	5.17%	1.92%	3.39%	8.45%	4.55%	0%	2.17%	2.04%	5.17%
Submissions failing on HWE	50%	3.85%	7.41%	12.00%	3.45%	7.69%	1.69%	7.04%	11.36%	1.67%	0%	6.12%	3.45%
Percentage of animal genotypes with No Nomination	3%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Submissions failing on excessive conflicts per chip	10%	57.69%	50.00%	22.00%	31.03%	19.23%	11.86%	18.31%	29.55%	28.33%	21.74%	26.53%	25.86%
Percentage of animal genotype reassigned		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%

CDCB
COUNCIL ON DAIRY CATTLE BREEDING

Check points during submissions

Genotype submission

- Sample Sheet file
- Final Report file

Check Program

Data QC file generated and deliver to the lab

- Lab check the genomic data based on the reports
  - If OK, upload the genotype
  - If not, correct the identified issue and re-start the process

Genotype uploading

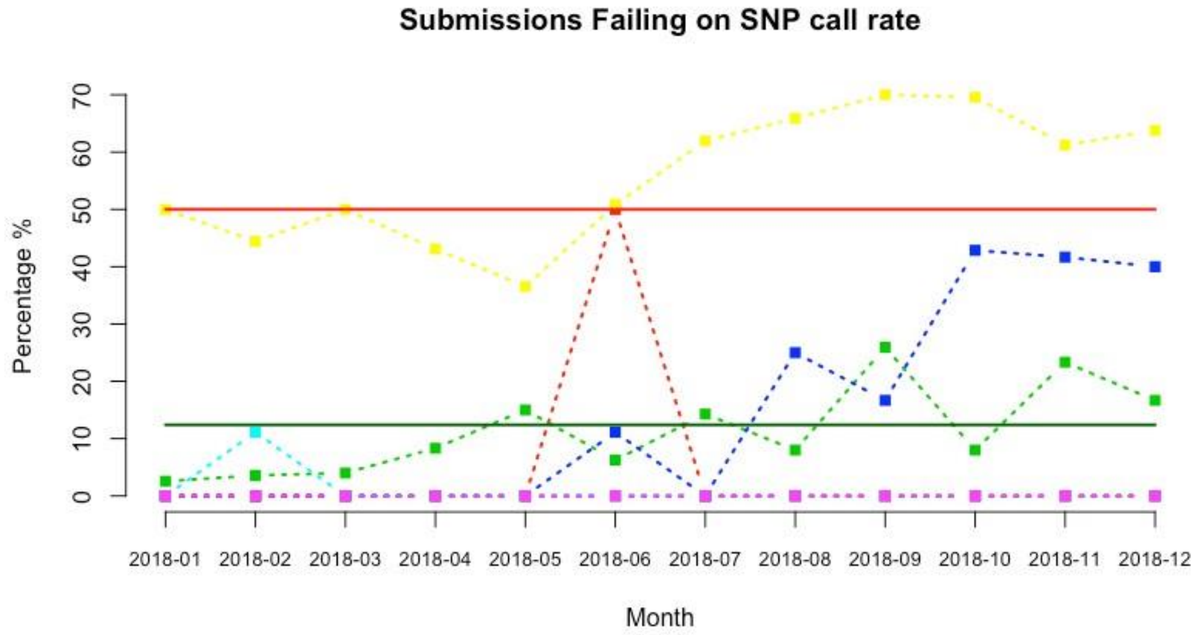
Data stored in the CDCB database &

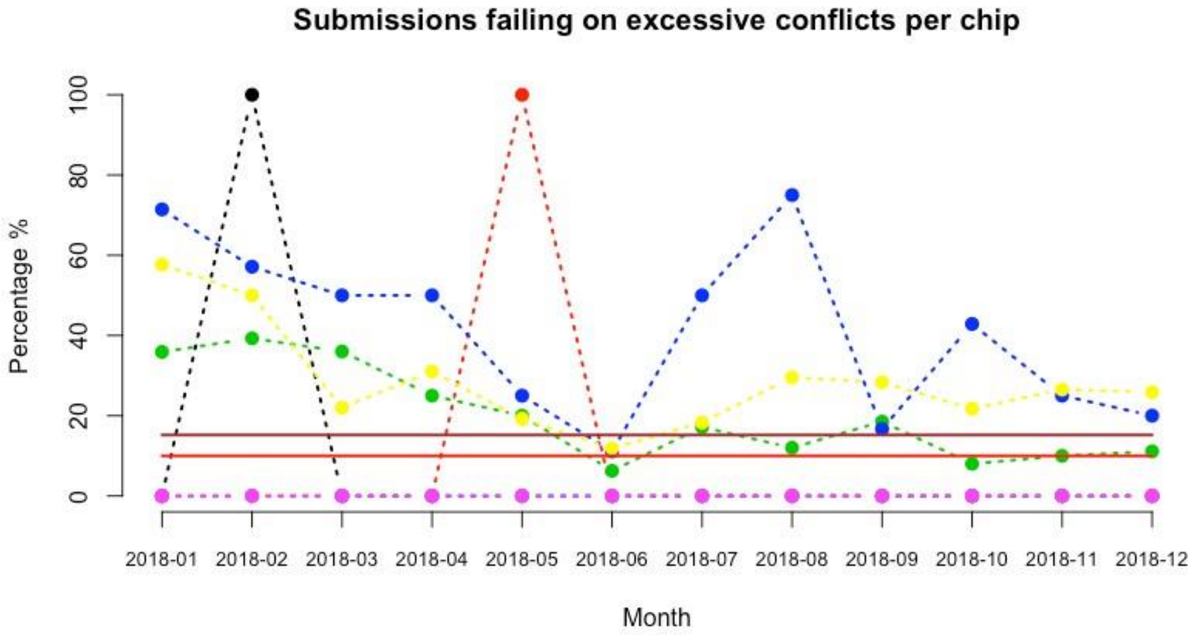
QC reports delivered to Lab and nominator

- **Nominators**
  - Genomic conflicts
  - Genomic errors
  - Parentage
  - Nominator report
- **Laboratories**
- Key not found
- No nomination
- No match sample
- Count.gt
- HWE
- Low call
- DataQC
- Genomic errors

CDCB
COUNCIL ON DAIRY CATTLE BREEDING

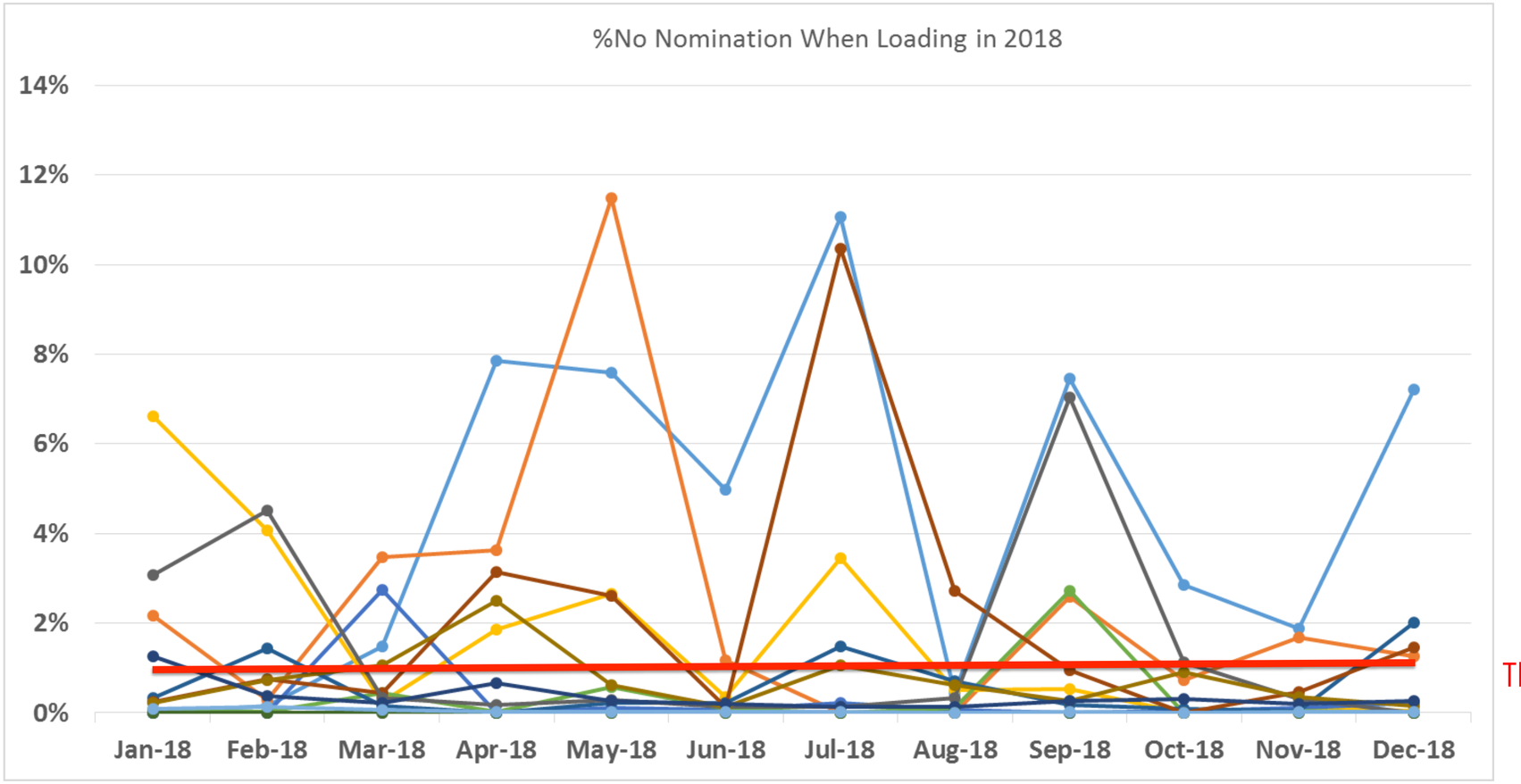
Performance metrics (Examples)





CDCB
COUNCIL ON DAIRY CATTLE BREEDING

Nominators' performance (Example)

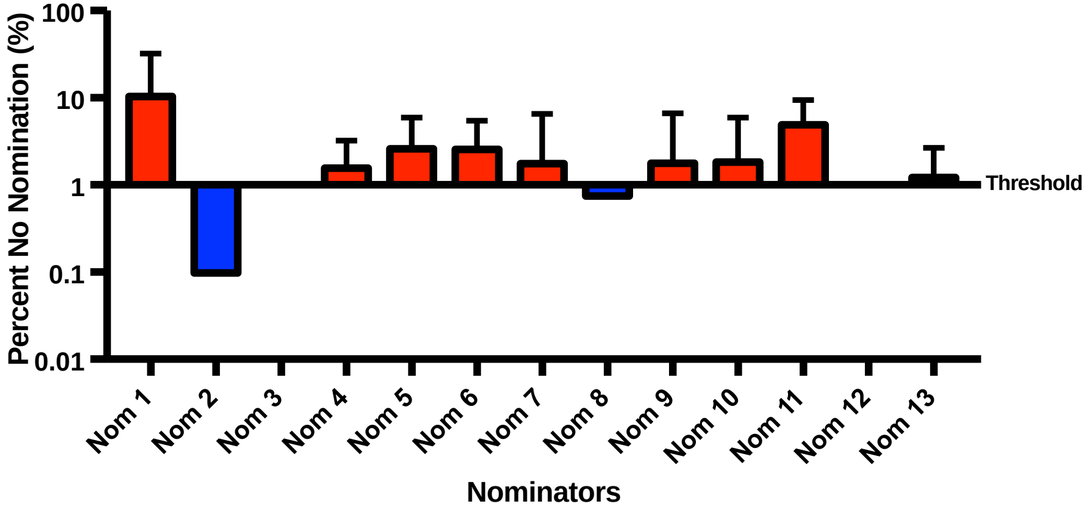


Threshold= 1%

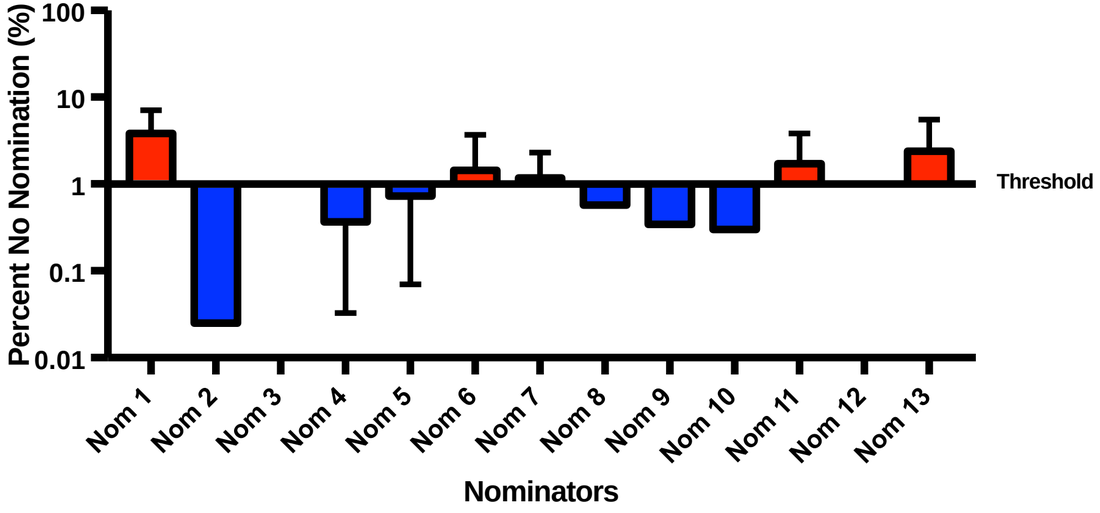
CDCB
COUNCIL ON DAIRY CATTLE BREEDING

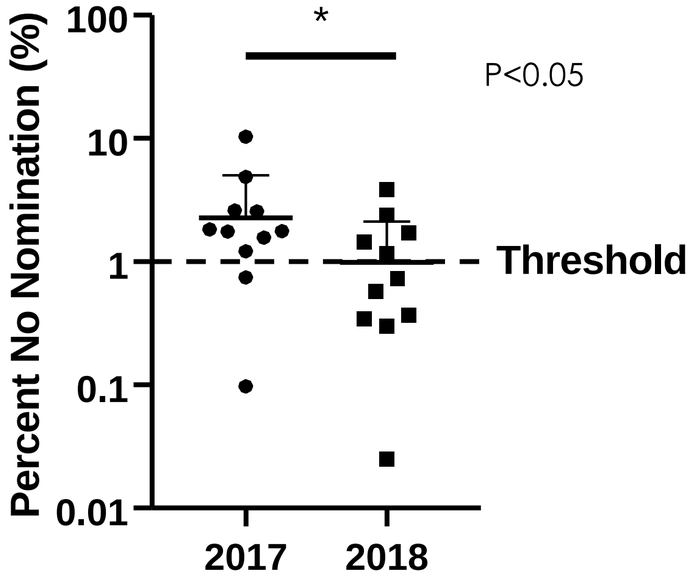
Impact of QC on Nominators' performance

2017



2018





CDCB
COUNCIL ON DAIRY CATTLE BREEDING

SUMMARY

- The CDCB has developed a customized QC system for evaluating laboratories performance
- Previous experience with the nominators demonstrated the positive impact after implementing a similar approach
- The QC program will assist the laboratories in delivering high quality data and contribute to maintaining the integrity of the CDCB database

ACKNOWLEDGEMENTS

- Participating dairy producers for supplying data
- DHI organizations and DRPCs for processing and relaying the information to the **Council on Dairy Cattle Breeding (CDCB)**
- Purebred breed associations for providing pedigree data

CDCB
COUNCIL ON DAIRY CATTLE BREEDING

Thanks for your attention

CDCB
COUNCIL ON DAIRY CATTLE BREEDING