

Roles and Responsibilities of Labs and Nominators

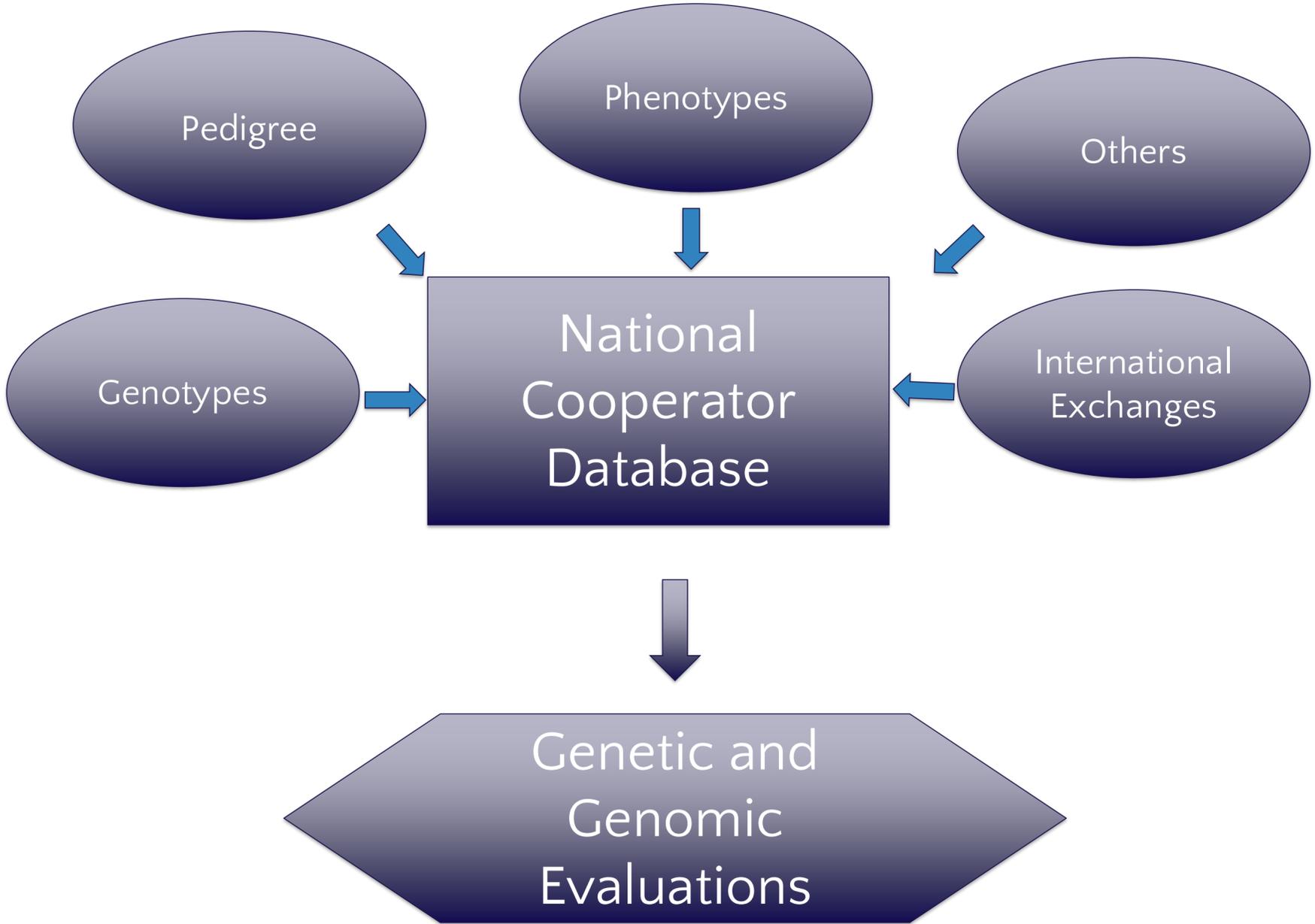
Kaori Tokuhisa

Genomic Data Analyst

7/30/2024

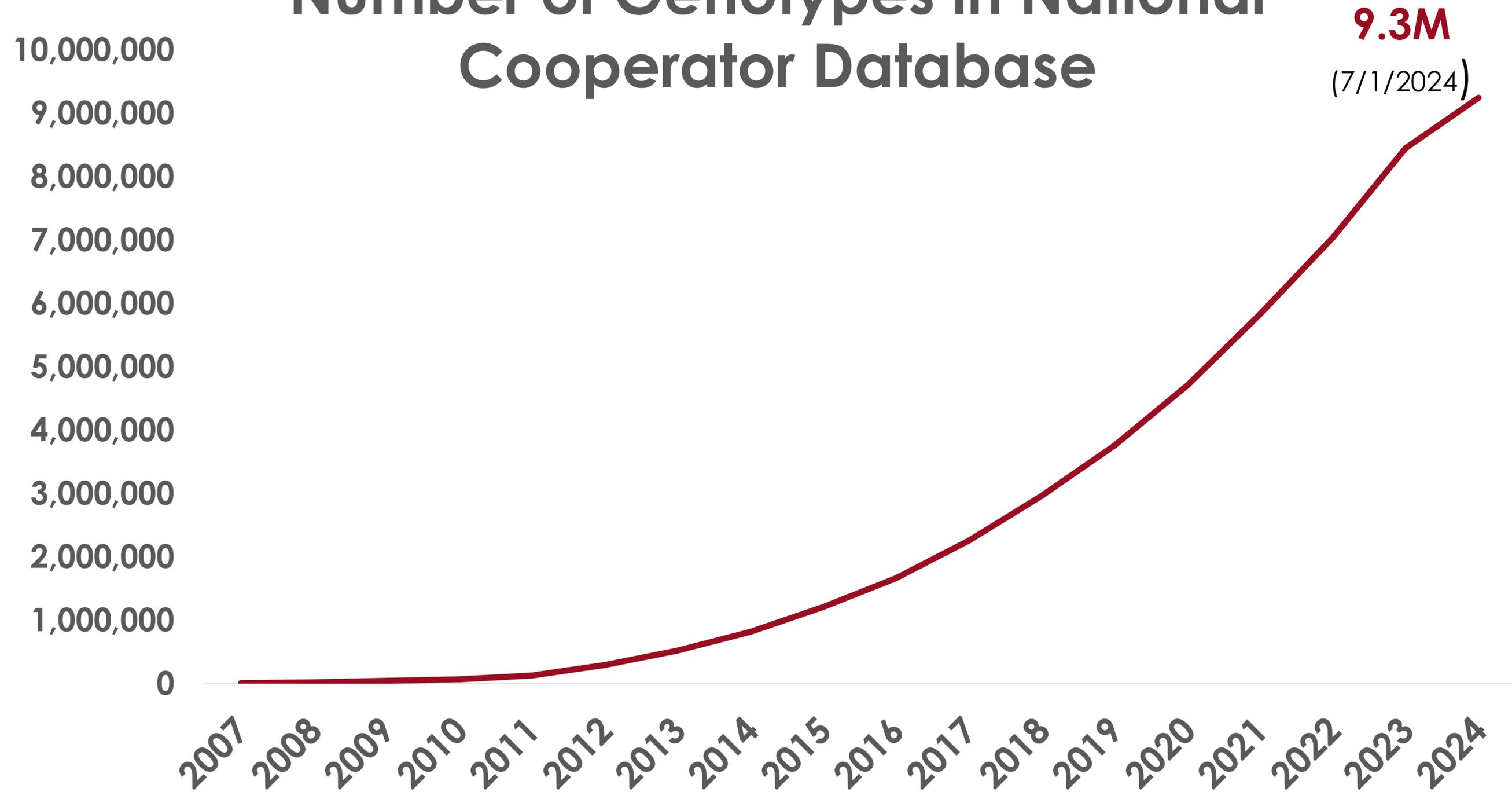


CDCB's Main Tasks



CDCB stewards the National Cooperator Database, distributes U.S. genetic evaluations, aggregates dairy cattle data, and conducts research.

Number of Genotypes in National Cooperator Database



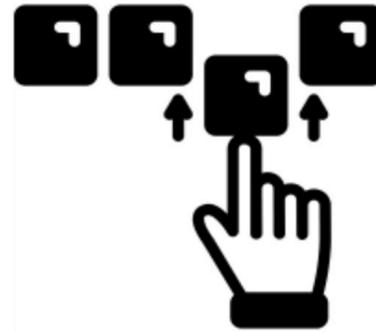
—Accum. GT

What is Data Quality?



Accuracy

usability code=N



Completeness

Sire/dam pedigree
missing

Relevancy

Relevancy

Genotype
Reassignment



Validity

CDCB blanked
dams due to
conflict



Timeliness

No nomination
when loading



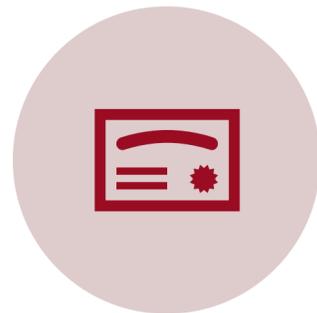
Consistency

Genotype
withdrawn

How Can CDCB Ensure the Data Quality?



WORKING WITH RELIABLE
DATA PROVIDERS



CERTIFICATION PROGRAM



EFFECTIVE QUALITY
CONTROL SYSTEM



GOOD COMMUNICATION
WITH DATA PROVIDERS



TRAINING

Core Requirements for Genomic Nominators



Collect DNA samples from the animal owner



Transfer DNA samples to a certified genotyping lab



Send pedigree and fee code; correct pedigree and genomic conflicts



Distribute CDCB results to the animal owner or marketing organization



Collect initial fees and forward to CDCB



Meet CDCB certification requirements

Core Requirements for Genomic Laboratories



Extract DNA from samples



Prepare single-nucleotide polymorphism (SNP) genotypes



Provide summary information back to the genomic nominator



Transfer the raw genotypes to the CDCB database



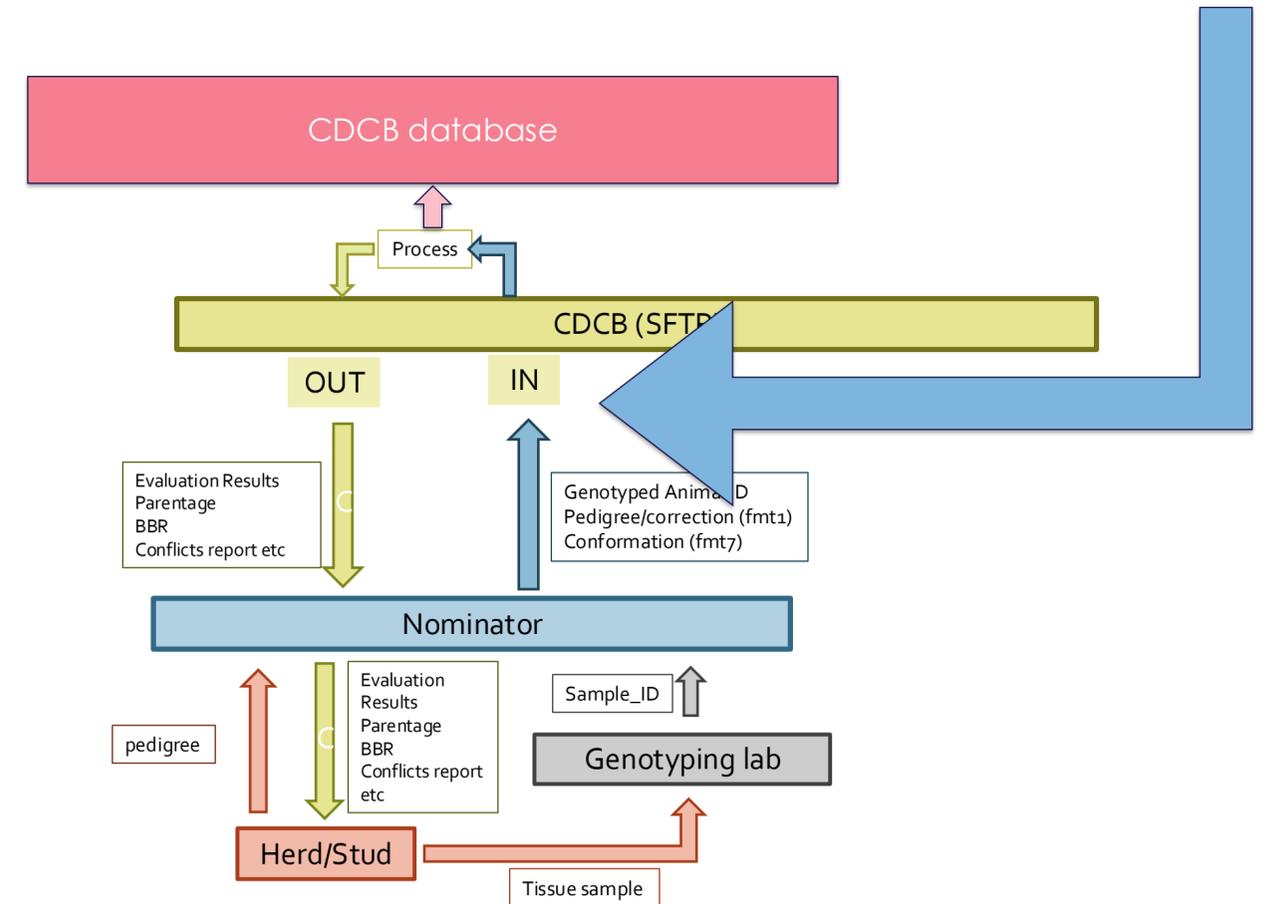
Meet CDCB quality certification standards

Nomination

YYYYMMDD.1GX

OFHOUSA000011111111HOUSA000000222222HOUSA000033333333A1B2C3D4E5F6 20160219B20160312G013HR000000NAMEOFANIMAL ET 35051162N2

- Nomination is a process of providing pedigree, assigning a fee code, and indicating what service you would like to receive.
- Nomination can be done either through SFTP (format1) or through WebConnect
- Nomination has to be done before the genotype submission
- Notify_VAL file will report nomination confirmation and errors



ANIMAL: BSUSA000068174286

1 Update Information 2 Review Changes and Submit

Nomination Status

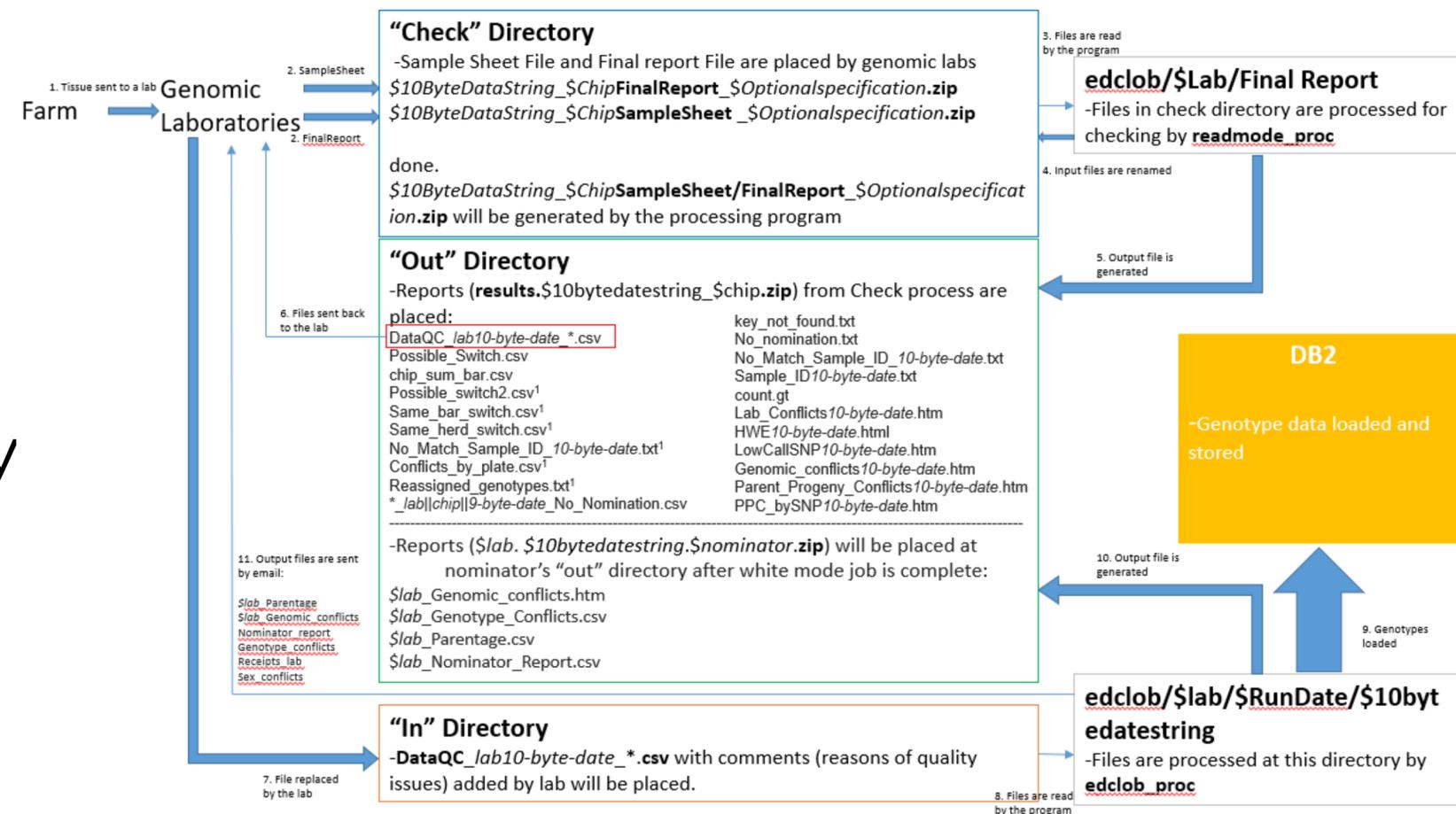
Nom. Date	Requester	Group/Herd	Herd Code Difference Reason Code	Fee	Fee Assigned Date	Mod Date	
2016-03-02	BS	35051162		2	2016-03-02	2016-03-12 12:00	
2024-06-04	ABS			1			

Cancel Next

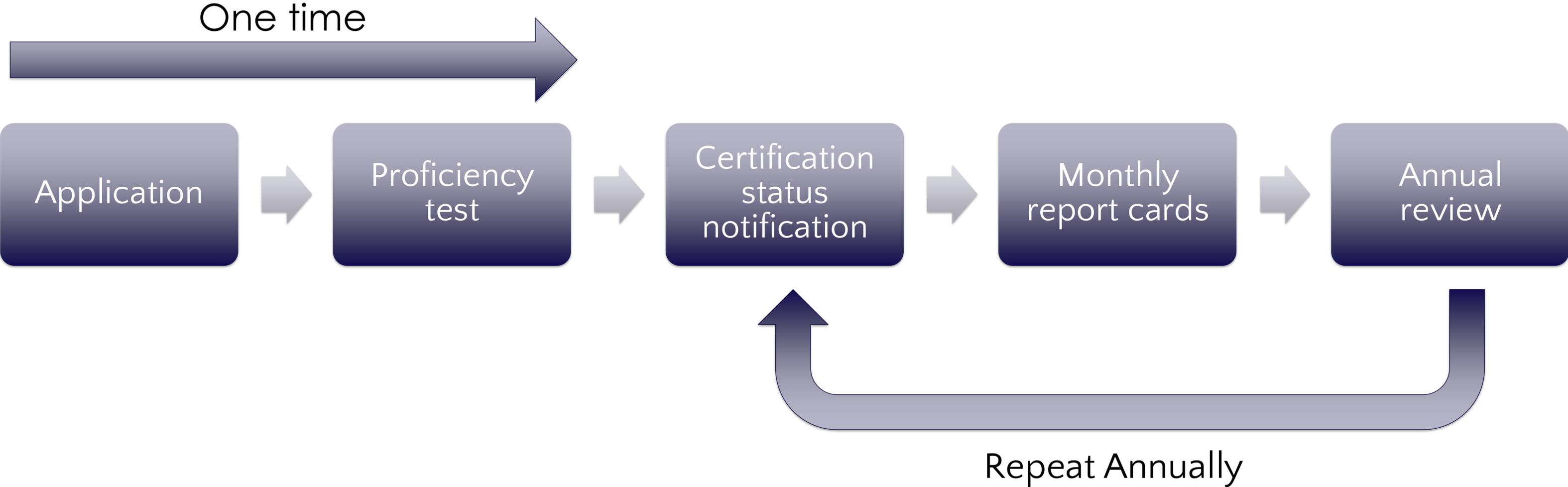
Genotype Loading

- Sample Sheet and Final Report files
- Loading process consist of 2 step (QC check+ Loading)

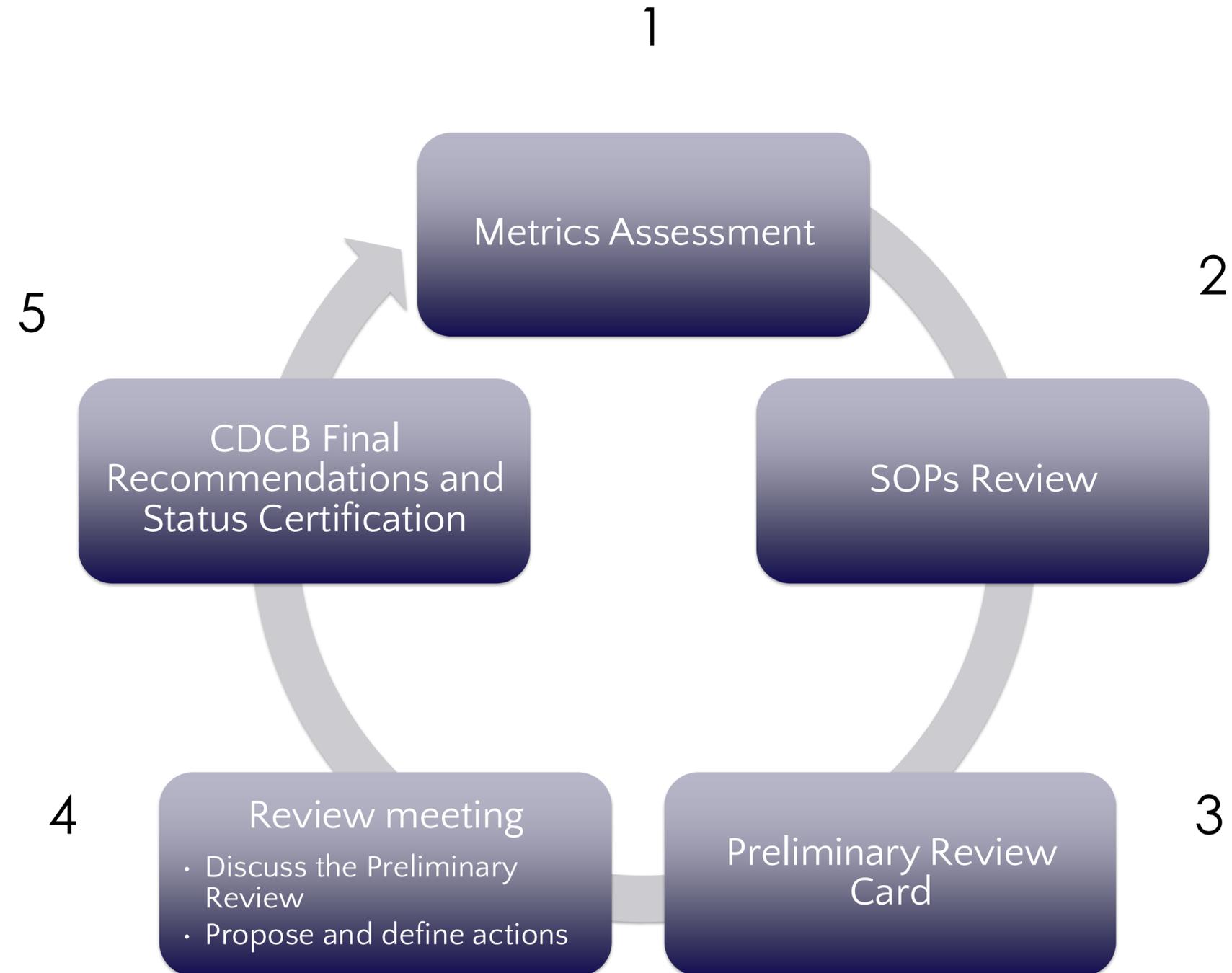
- QC files are generated from each step
- Data QC file placement in “in” directory triggers loading process



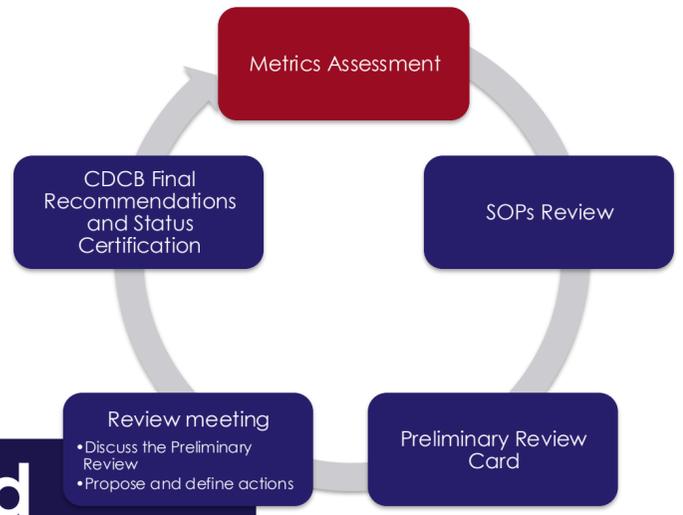
CDCB - Certification Process



Annual Review Plan

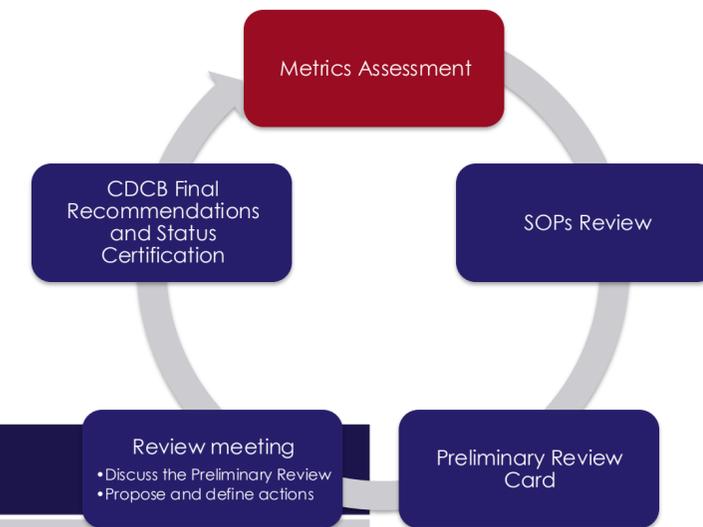


Current Nominator QC metrics



Class	Metric	Threshold
Critical	No Nomination when loading	1%
Critical	Unknown Animal ID	1%
Major	CDCB blanked dam due to conflict	2%
Major	Usability=N	5%
Major	Fee Code=N	1%
Major	Genotype withdrawn	1%
Major	Genotype reassigned	1%
Minor	Changes in pedigree	30%
Minor	Sire pedigree missing	1%
Minor	DamID or pedigree missing	25%
Info	Herd reason code	NA
Info	Nomination Fee Codes changed	NA
Special	Any of above metrics	Percentage Exceeds Average of Previous 3 Months by > = 5%

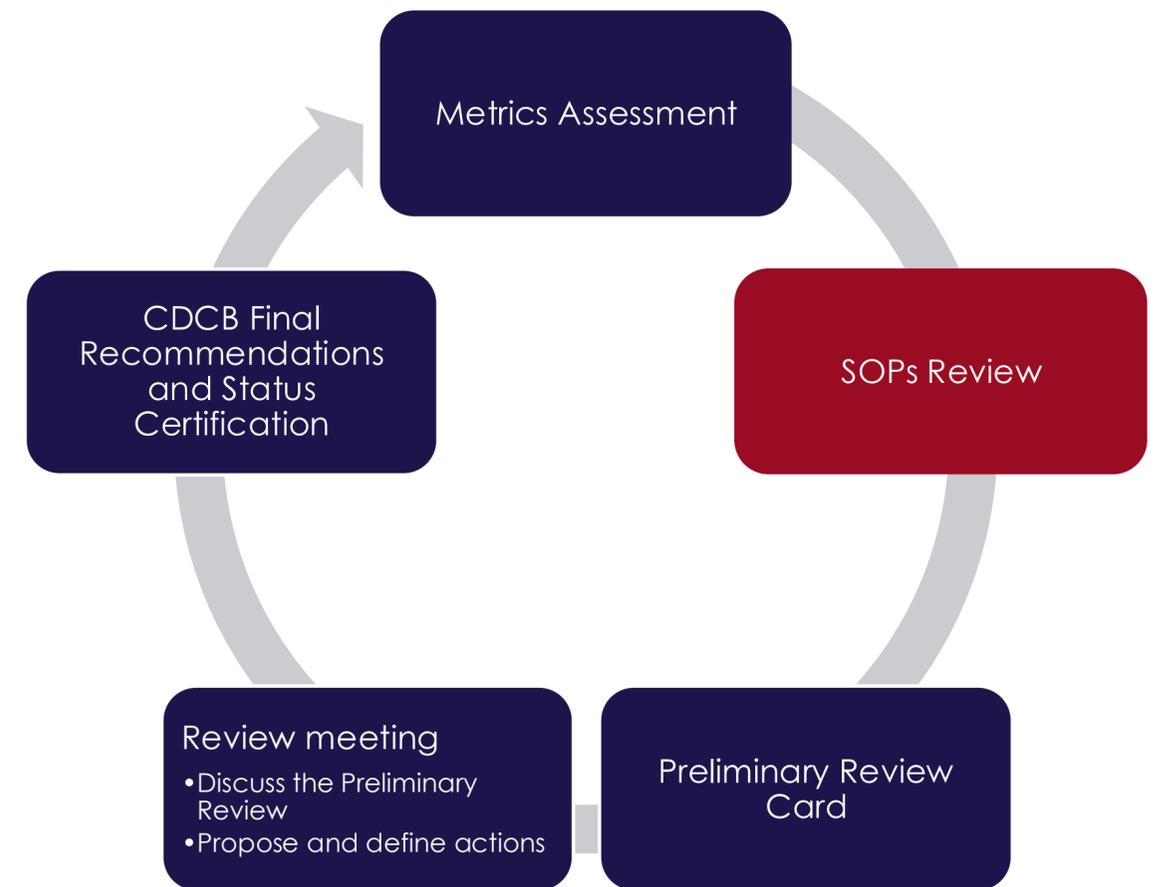
Current Lab QC metrics



Class	Metric	Threshold
Critical	Submission with <10 animal genotypes	10%
Critical	Submission failing on SNP call rate (across animal)	50% on submissions with 50 or more genotypes.
Critical	Submissions failing on SNP PPC	25% on submissions with 50 or more genotypes.
Critical	Submissions failing on HWE	50% on submissions with 50 or more genotypes
Info	Percentage of animal genotypes with No Nomination	NA
Info	Submissions flagged on excessive conflicts per chip	NA

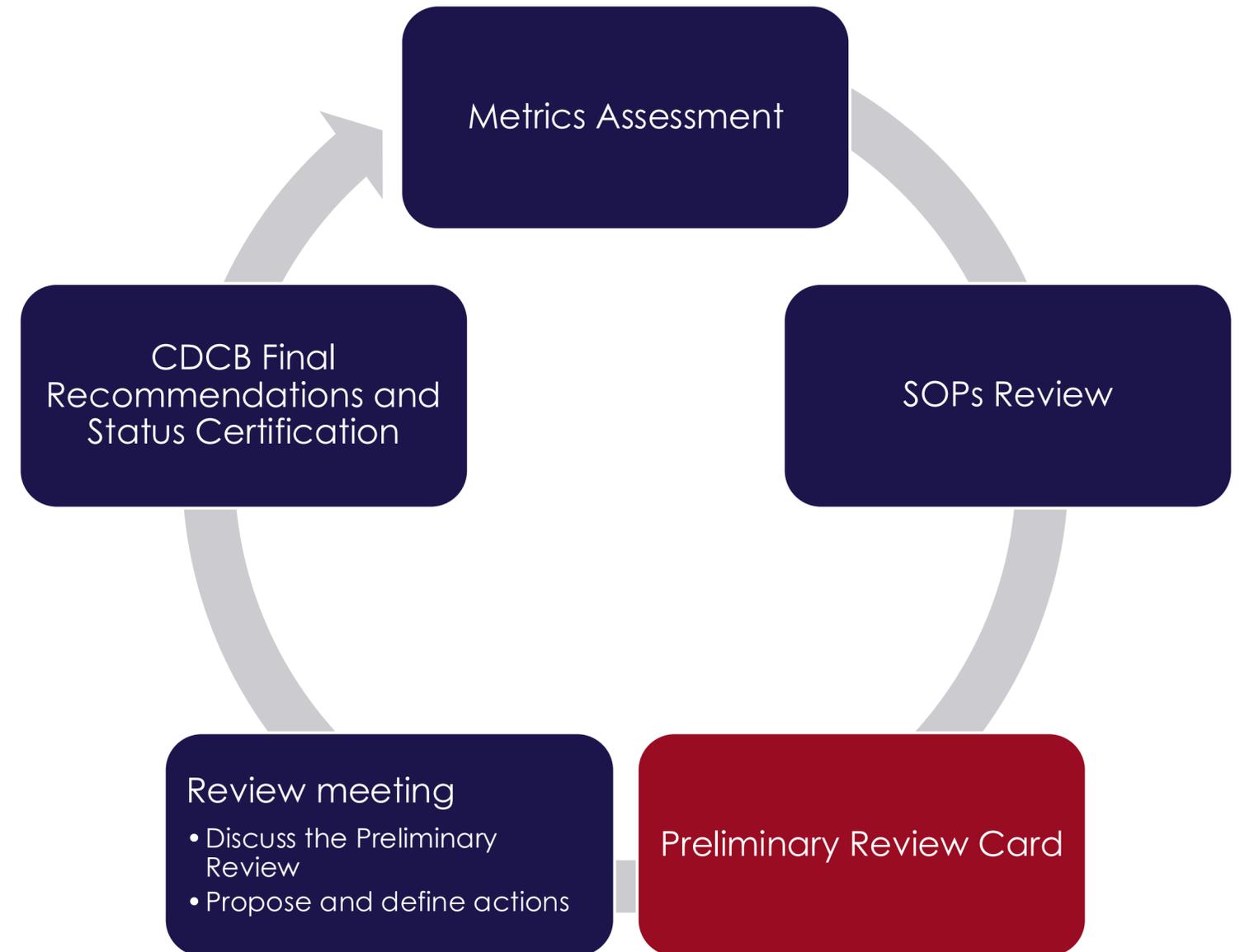
2. SOP Review

- Should be updated every year as new changes or improvements are made in the procedure
- Our team compares the updated SOPs with the previous year's to see if you provided sufficient descriptions in your SOP or if our previous suggestions or newly implemented procedures are incorporated in the latest SOP.
- Protocols within the SOPs should be replicable; you should describe the processes in enough detail that a new employee can follow them without trouble.
- Iteration of 2-3 years to be mature.
(Most organizations have been working with CDCB for > 2 years)



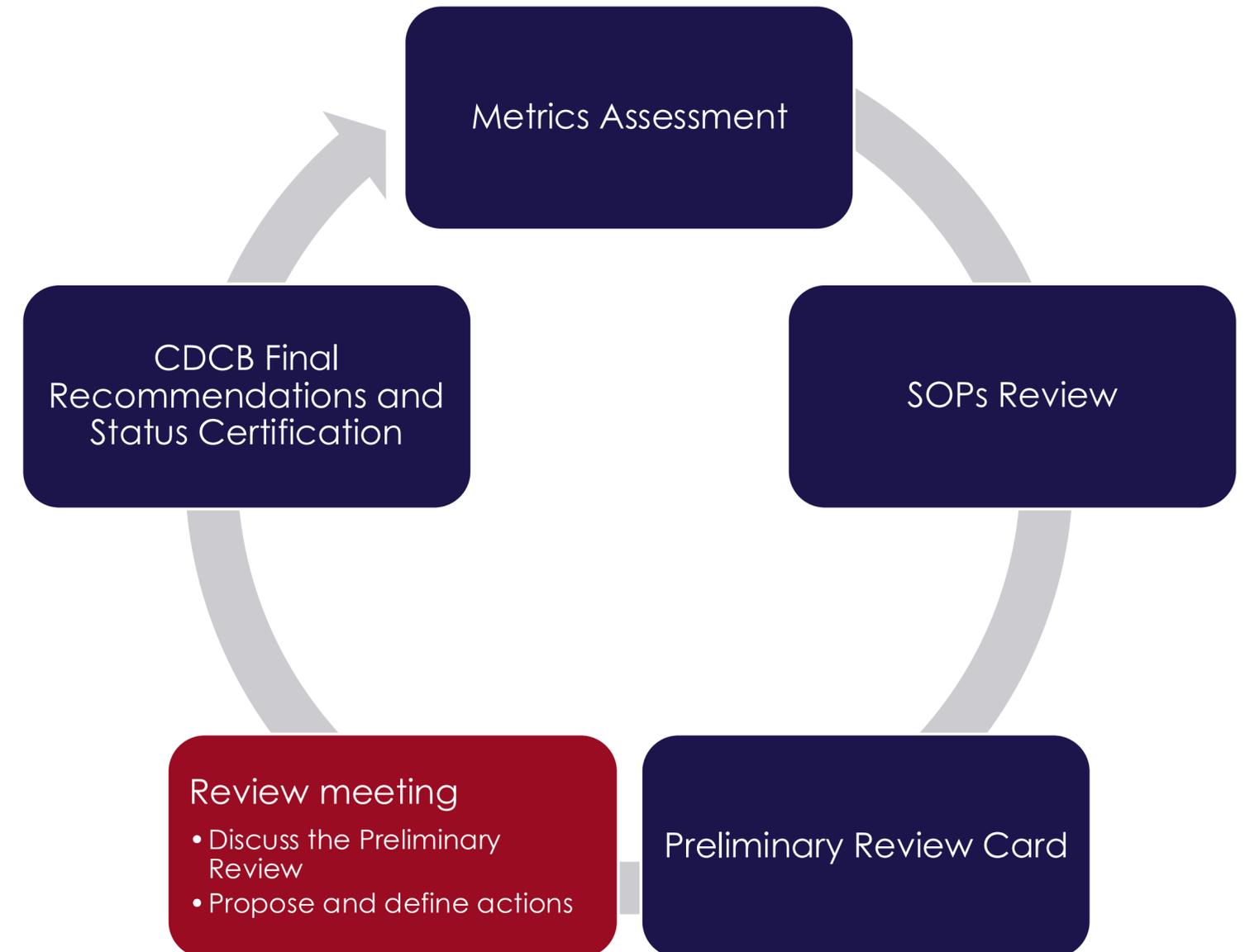
3. Preliminary Review Card

- This is a summary of the performance metrics, SOP, and any other observations that we had for the previous year as a preparation for the conference call
- We send preliminary report card in advance, so you have enough time to review and prepare for the meeting.



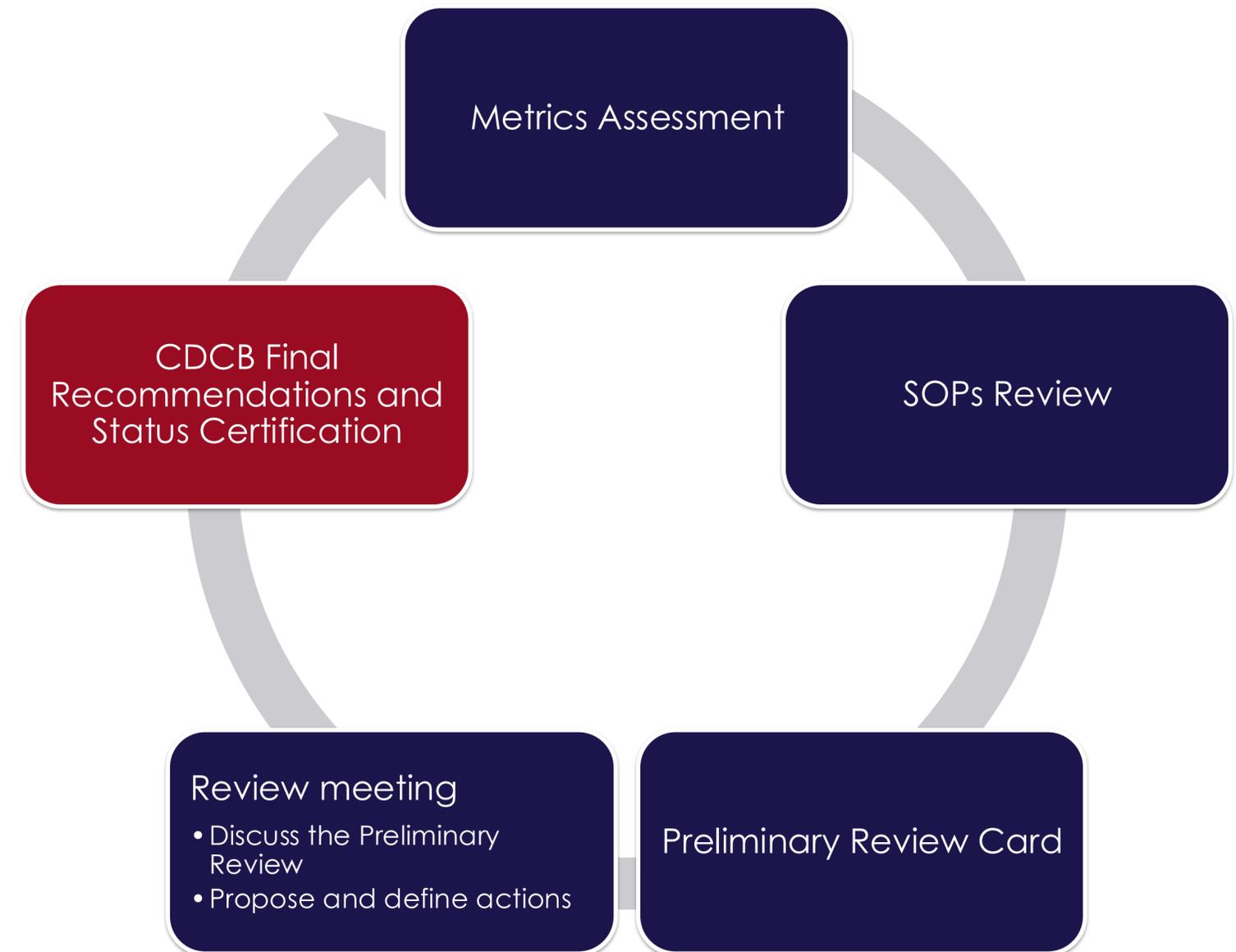
4. Review Meeting

- 45 mins conference call with each nominator and lab.
- We present each performance and SOPs, observations and comments
- We discuss issues and provide suggestions/advices to resolve them and set a plan for the coming year.
- Time to exchange information from both side



5. CDCB Final Recommendations and Status Certification

- **Final report card** is a summary of all info included in the preliminary report card and topics discussed during review meeting.
- The **final report card** also includes certification status for the coming year (*conditional, provisional, certified, decertified*) and the justification.



Summary

- The roles and responsibilities of labs and nominators have been more important due to the **increasing volume** and **complexities of data** entering CDCB system.
- Reliability and accuracy of CDCB evaluations are based on **systematic check** and **collaborator's effort** on ensuring the quality of data
- To ensure certain level of proficiency, CDCB requires collaborators to comply with CDCB requirements
- Collaborative effort among nominators, labs, and CDCB is key to generate accurate evaluations and dairy genetic improvements.

Thank you!!

