Implementation of adding discovered grandsires and great grandsires using constructed ID

José A. Carrillo[1], Lillian R. Bacheller[1], Jay Megonigal[1], Ezequiel L. Nicolazzi[1], Paul VanRaden[2] and George R. Wiggans[1]

[1]CDCB - Council on Dairy Cattle Breeding, Bowie, MD, USA

[2]Animal Genomics and Improvement Laboratory, ARS, USDA, Beltsville, MD, USA

Interbull Meeting - Lyon, France

August 26, 2023

CDCB

COUNCIL ON DAIRY CATTLE BREEDING

Pedigree validation in the US database

Over 7 million animals genotyped in U.S. system

The pedigree of each animal entering the database is checked (counting opposite homozygotes), SNP at a time

Each genotype compared with most others to discover identical genotypes and parent-progeny relationships

All conflicts resolved

Genotypes imputed to 79K using **Findhap.f90,** which provides haplotypes

Portion of validated parents of genotyped animals born in 2022 (98% of sires, 57% of dams)

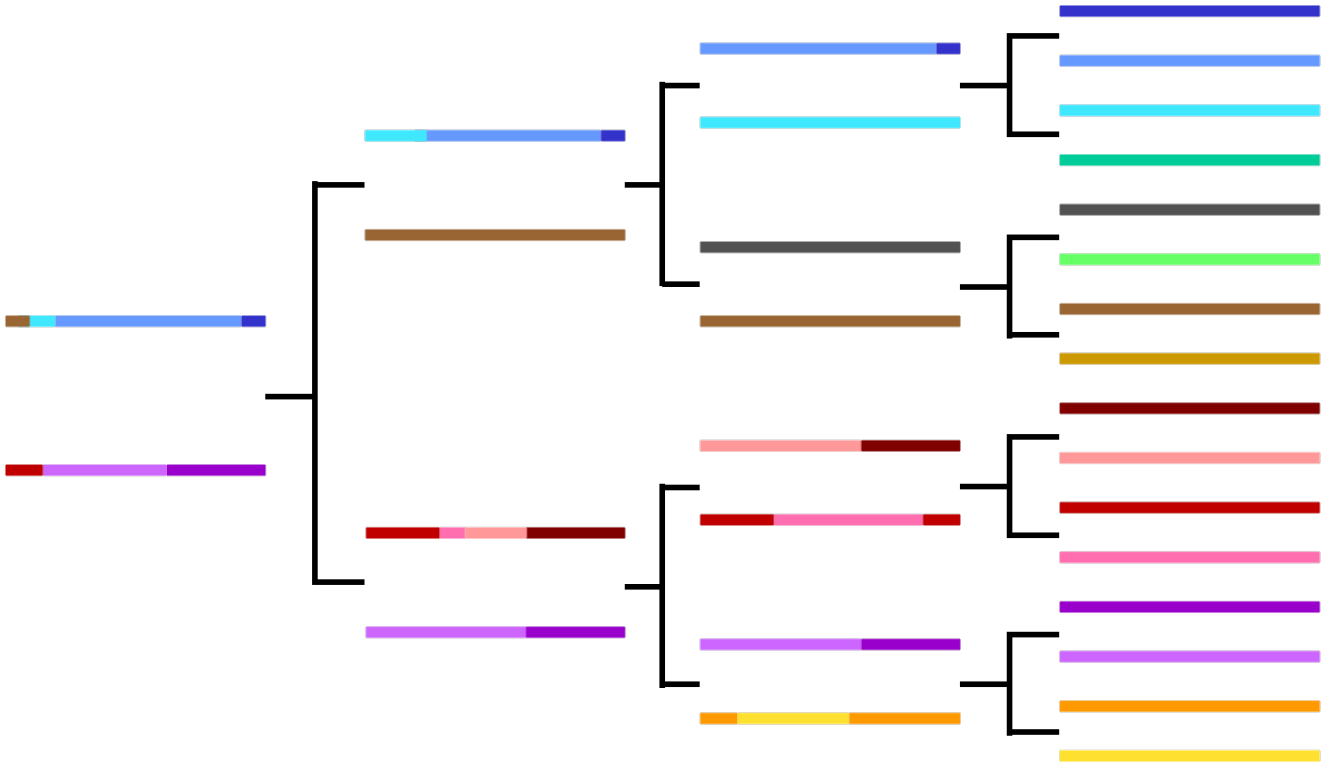
CDCB
COUNCIL ON DAIRY CATTLE BREEDING

USDA

Discovering grandsires and great grandsires

Fixped.f90 uses haplotypes to discover distant relationships as MGS and MGGS

(VanRaden et al., 2013)

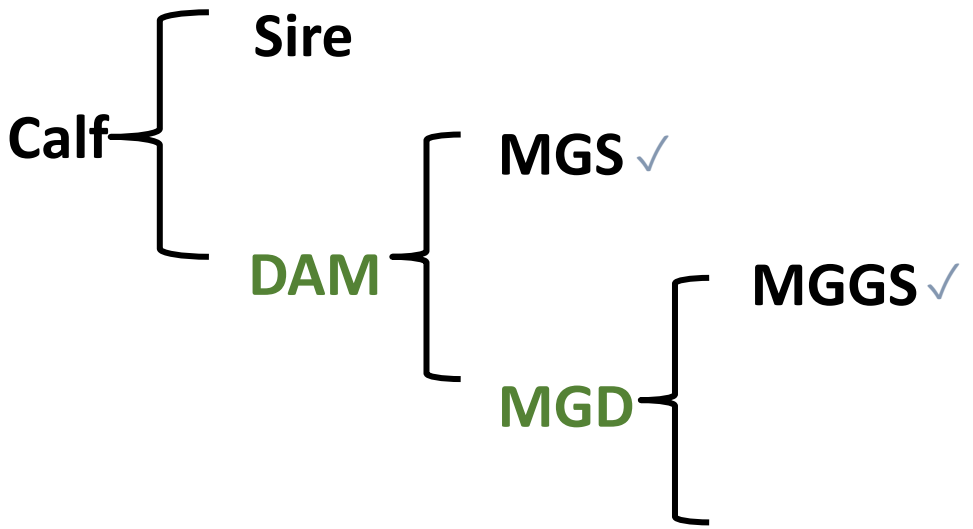


https://aipl.arsusda.gov/publish/jds/2013/96_1874.pdf

Linking discovered MGS and MGGS

Finddam creates the **constructed dam** and/or **MGD ID** to link calves to MGS and MGGS in the pedigree

- ~ 1.79 million Constructed ID will be created according to the Monthly August 2023 evaluation
- 67K already created and added to the pedigree
- Total of 1.86 M Constructed ID



CDCB
COUNCIL ON DAIRY CATTLE BREEDING

USDA

Filling in pedigrees

- First check if true dam can be discovered in same herd
  - Match birth and fresh dates, only 1 dam's pedigree matches calf's (only non-ambiguous cases are filled)
- Discovery of registered animals are reported to breed associations only

Already discovered and added 400,000 MGS and MGGS whose dam ID and MGD ID were previously reported by the owners

CDCB
COUNCIL ON DAIRY CATTLE BREEDING

USDA

Filling in pedigrees

- **Constructed ID**: **HO USA DAM (MGD)** calf internal numeric ID (key number)

HOUSADAM004235395

HOUSAMGD004235395

- **For international IDs,** the only difference will be the 3-letter breed code (HOL instead of HO) and the addition of the sex (F or M) after the country code

HOLUSAFDAM004235395

- **To further facilitate the constructed ID recognition,** the name of each constructed animal will be **"Dam of [ID of source animal]"** or "**MGD of [ID of source animal]"**

CDCB
COUNCIL ON DAIRY CATTLE BREEDING

USDA

Constructed IDs standards

- **Unique:** must be unique in the pedigree, as they are only used to link source animals to their ancestors
- **Traceable**: must be connected to the source animal from which it was derived (and the country that generated it).
- **Stable:** The connection between the constructed ID and its source animal must be perpetual – unless the true ancestor is found.
- **Recognizable:** must be easily recognizable as placeholders and never be considered as the ID of a true ancestor.

CDCB
COUNCIL ON DAIRY CATTLE BREEDING

USDA

Constructed IDs rules

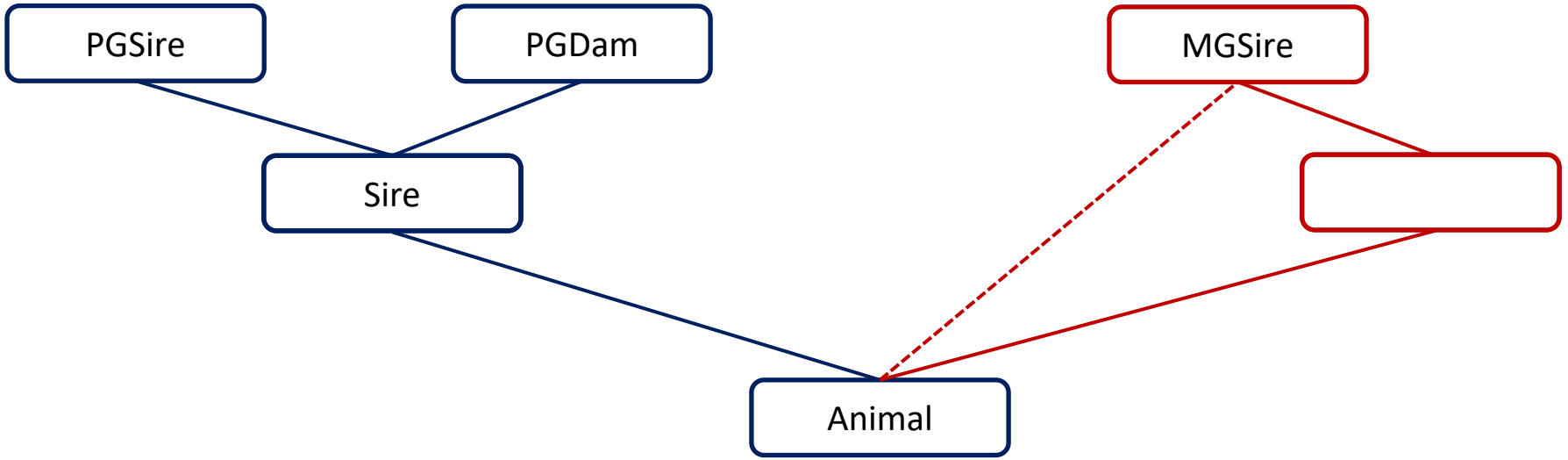
- **Constructed IDs** will **ONLY** be generated by the CDCB
- Other submitters can modify the pedigree record as usual
- Submitters will **ONLY** be allowed to replace constructed IDs with a true ancestor ID or delete the constructed ID to reject the connection to the MGS or the MGGS
- Users are **NOT** allowed to replace constructed IDs with alternative constructed IDs

CDCB
COUNCIL ON DAIRY CATTLE BREEDING

USDA

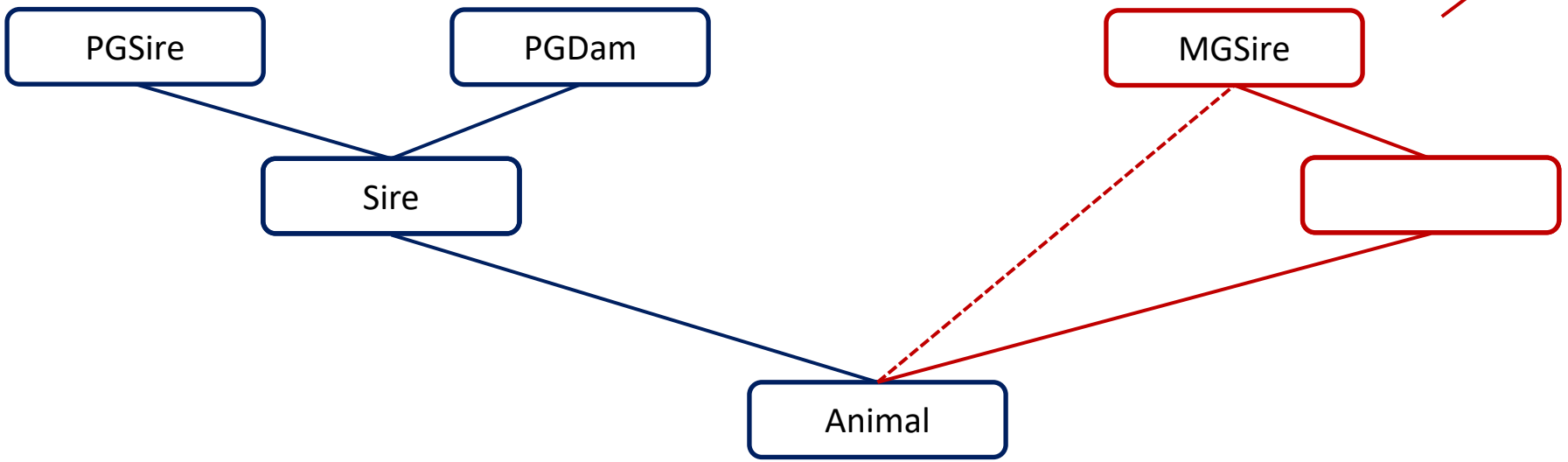
Dam unknown; Dam suggested based on herd, sire, and calving date

HO840000013897236
HO840000013897237
HO840000013897238
HO840000013897239
HO840000013897240

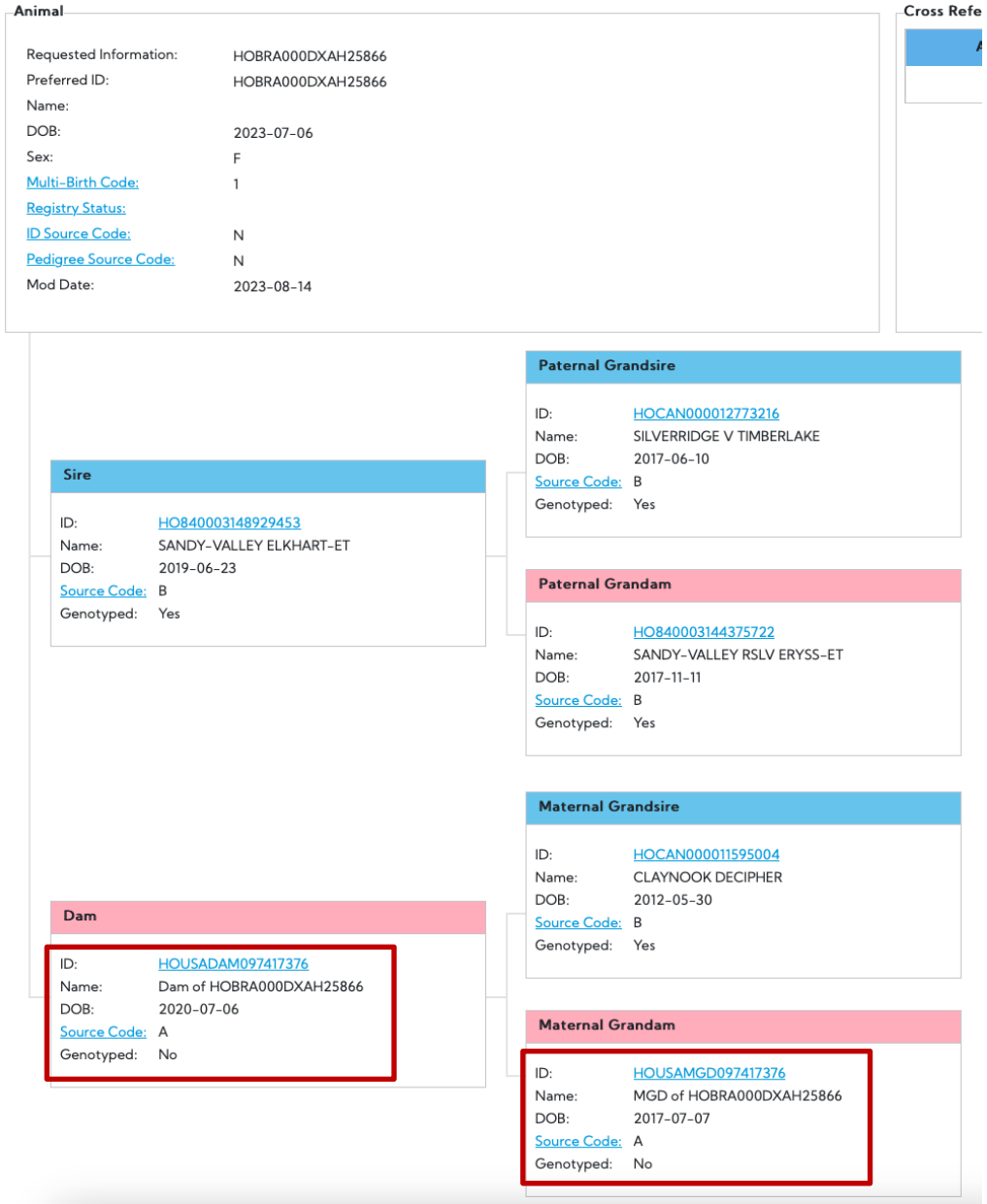


Dam unknown; Dam ID constructed, and pedigree created with discovered MGS as sire

HO840000013897236
HO840000013897237
HO840000013897238
HO840000013897239
HO840000013897240



Real example from WebConnect



Animal

Requested Information:	HOBRA000DXAH25866
Preferred ID:	HOBRA000DXAH25866
Name:	
DOB:	2023-07-06
Sex:	F
Multi-Birth Code:	1
Registry Status:	
ID Source Code:	N
Pedigree Source Code:	N
Mod Date:	2023-08-14

Cross Refe

Sire

ID:	HO840003148929453
Name:	SANDY-VALLEY ELKHART-ET
DOB:	2019-06-23
Source Code:	B
Genotyped:	Yes

Paternal Grandsire

ID:	HOCAN000012773216
Name:	SILVERRIDGE V TIMBERLAKE
DOB:	2017-06-10
Source Code:	B
Genotyped:	Yes

Paternal Grandam

ID:	HO840003144375722
Name:	SANDY-VALLEY RSLV ERYSS-ET
DOB:	2017-11-11
Source Code:	B
Genotyped:	Yes

Dam

ID:	HOUSADAM097417376
Name:	Dam of HOBRA000DXAH25866
DOB:	2020-07-06
Source Code:	A
Genotyped:	No

Maternal Grandsire

ID:	HOCAN000011595004
Name:	CLAYNOOK DECIPHER
DOB:	2012-05-30
Source Code:	B
Genotyped:	Yes

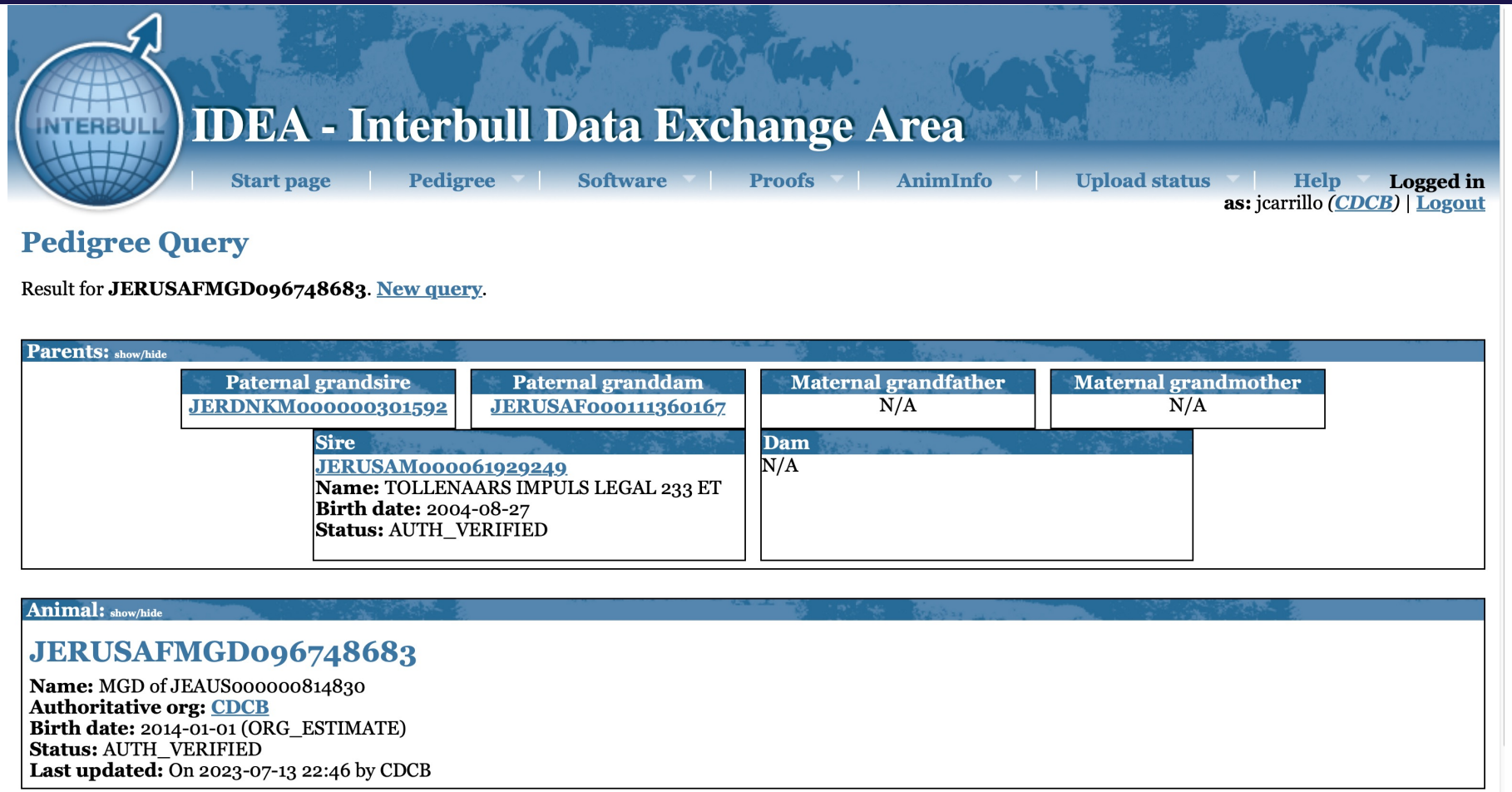
Maternal Grandam

ID:	HOUSAMGD097417376
Name:	MGD of HOBRA000DXAH25866
DOB:	2017-07-07
Source Code:	A
Genotyped:	No

CDCB
COUNCIL ON DAIRY CATTLE BREEDING

USDA

Real example from IDEA



Filled pedigree results

J. Dairy Sci. 103
https://doi.org/10.3168/jds.2019-17580

Discovering ancestors and connecting relatives in large genomic databases

J. P. Nani,[1,2*] L. R. Bacheller,[3] J. B. Cole,[1] and P. M. VanRaden[1]
[1]USDA, Agricultural Research Service, Animal Genomics and Improvement Laboratory, Beltsville, MD 20705-2350
[2]Estación Experimental Agropecuaria Rafaela, Instituto Nacional de Tecnología Agropecuaria, 22-2300, Rafaela, SF, Argentina
[3]Council on Dairy Cattle Breeding, Bowie, MD 20716

Table 2. Traditional and genomic EBV value means, SD, and reliabilities for yield traits of 295,136 animals with newly found ancestors, before and after pedigree completion

		Incomplete pedigree			Complete pedigree		
EBV	Trait	Mean (kg)	SD (kg)	Reliability (%)	Mean (kg)	SD (kg)	Reliability (%)
Traditional	Milk	1,948	720	26.6	2,064	811	32.6
	Fat	72.2	15.1	25.0	76.7	18.2	32.0
	Protein	59.4	14.6	26.9	63.0	17.0	32.9
Genomic	Milk	2,186	492	76.2	2,258	513	77.1
	Fat	74.9	19.7	76.0	77.6	20.5	76.9
	Protein	63.6	13.4	76.3	65.7	14.1	77.3

Impact on evaluations

- CDCB performed a full test-run (traditional and genomics) in November 2022 to assess the impact of the **full** implementation of discovered pedigrees
  - This will NOT happen, as there will be a gradual implementation throughout 2023
- AI bulls: **Nearly no impact**, some marginal variation on some bulls adding a lot of daughters.
- Animals adding pedigrees: variable impact depending on the entity of pedigree changes and discovered MGS/MGGS. Large increase in reliability (e.g., better predictions!)
- Results in line with Nani et al., even if the number of inclusions was more than twice

CDCB
COUNCIL ON DAIRY CATTLE BREEDING

USDA

Summary and implementation

- Linking disconnected animals improved genetic and genomic estimations
- Correcting pedigree errors generate more accurate inbreeding estimates
- Real dam IDs are preferred over constructed IDs
- Pedigree providers have the option to remove discovered relationships
- Constructed ID implementation started on April 2023
- Addition of historical data is still in process
- An increase of 1 point in average reliability has been confirmed on much larger population

CDCB
COUNCIL ON DAIRY CATTLE BREEDING

USDA

Acknowledgements and disclaimers

- 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

- Mention of trade names or commercial products is solely for the purpose of providing specific information and does not imply recommendation or endorsement by CDCB
- CDCB is an equal opportunity provider and employer

CDCB
COUNCIL ON DAIRY CATTLE BREEDING

USDA

Thank you!