Extending ICAR parentage verification certification to whole-genome SNP comparisons

E.L. Nicolazzi[1*], M. McClure[2], G. Jansen[1], and G.R. Wiggans[1]

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

[2] *ABS Global, DeForest, WI, USA*

Current ICAR/ISAG PV SET approach

- **ISAG PV SET composition**
  - Standard panel of 195 SNPS (recently revised) used internationally for parentage verification.
- **Quality requirements**
  - Requires ≥ 95% call rate on the PV SET (at least 185 successfully called SNPs)
  - Minimum of 175 homozygous common SNP comparisons per test required
- **Classification thresholds**
  - Fixed thresholds specifically tuned to the 195 SNP set (200 original)
  - Duo and trio comparisons

Why whole-genome SNP comparisons?

- **Exclusions due to fixed-panel approach**
  - Many platforms do not guarantee coverage of all 195 ISAG PV SNPs, resulting in “LOWCALL” failures
  - Re-genotyping is often impossible due to unavailability of biological samples
- **Heterogenous genotyping ecosystem**
  - Landscape includes 60+ arrays ranging from 3K to 777K SNPs
  - Sequencing-based technologies growing
- **Need for platform-agnostic standards**
  - An alternative (or complementary) method leveraging all available high-quality SNPs is required

Objectives

- **Assess the situation**
- **Establish empirical thresholds on a whole-genome PV situation**
  - Empirically derived
  - Easy-to-apply thresholds for duo-based, whole-genome SNP comparisons
  - Robust across genotyping platforms
- [desired] **Maintain ICAR classification logic**
  - Mirror the existing ACCEPTED / DOUBTFUL / EXCLUDED logic
- **Validate performance**
  - Quantify performance across breeds and densities
  - Assess the value of trio-based verification in high-density contexts

Materials and methods

Data Source and Quality Control

- U.S. National Cooperator Database
- Autosomal markers
- Pre-exclusion of genotypes with < 90% call rate
- Removed clones, split embryos, and identical twins

Genotyping Platforms

- 56 certified SNP arrays ranging from ~3K to 777K SNPs (currently 60+, including WGS-based systems)
- Diverse technologies representing different densities and manufacturers

Testing Populations

- Balanced representation across 5 breeds: Holstein, Jersey, Brown Swiss, Guernsey, Ayrshire
- ~100,000 DUO pairs: ***Known*** true parent-offspring relationships (max 5 per animal)
- ~100,000 NOT DUO pairs: Unrelated pairs + controlled set of full/half sibs (~10% each)

Analysis Pipeline

- 195 vs. WG performance
- Mendelian conflicts per pair
- Metric: % conflicts over total informative homozygous comparisons

Performance
ISAG PV SET

METRIC	ISAG PV SET (195 SNPS)
Exclusions (LOWCALL)	~25% Excluded 42,758 comparisons failed call rate
DUO Ambiguity	754 Doubtful 1,367 Rejected
NOT DUO False Matches	1,865 Wrongly Accepted Large false positive risk
NOT DUO Ambiguity	5,765 Doubtful

Confirmed: LOWCALL is a concern in the existing commercial landscape

Performance: While the 195 performs relatively well, there are still relatively high wrong calls and doubtful results

Performance and thresholds of WG SNPs

SENSITIVITY (TPR)

99.83%

On declared DUO pairs

SPECIFICITY (TNR)

99.97%

On declared NOT DUO pairs

Key Empirical Findings

- **0.6% Acceptance Threshold:** Selected because virtually no unrelated animals fell below this value (except 2 likely pedigree errors).
- **1.0% Rejection Threshold:** Provides clear separation; only a small number of full-sibs were misclassified as "Doubtful" rather than Rejected.
- **The "Doubtful" Zone (0.6% – 1.0%):** Effectively captures edge cases, often linked to low-density chips (e.g., 3K) or specific array quality issues.

Performance Comparison
ISAG PV SET vs. Whole-genome

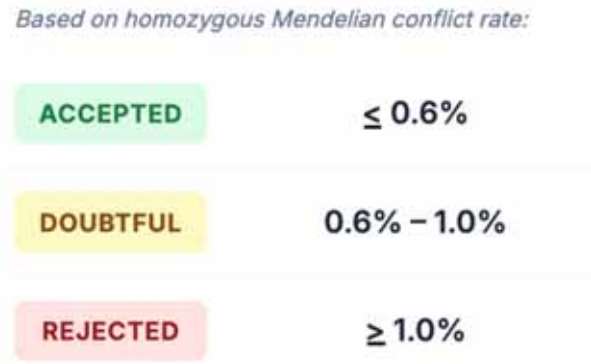
METRIC	ISAG PV SET (195 SNPS)	WHOLE-GENOME (PROPOSED)
Exclusions (LOWCALL)	~25% Excluded 42,758 comparisons failed call rate	Minimal Pre-selected for 90% call rate
DUO Ambiguity	754 Doubtful 1,367 Rejected	161 Doubtful Reduced to 19 with QC SNPs
NOT DUO False Matches	1,865 Wrongly Accepted Large false positive risk	2 Wrongly Accepted High specificity
NOT DUO Ambiguity	5,765 Doubtful	26 Doubtful Reduced to 12 with QC SNPs

Duo comparisons:
Highly discriminative on their own

Trio comparisons:
No benefit

Conclusions and recommendations

- When available, whole-genome comparisons preferred over ISAG PV SET in dairy cattle
  - *Virtually* eliminates "LOWCALL" loss
  - Only use on animals with +7K SNP density or SNP arrays without known QC issues
    - Watch for chromosomal abnormalities!
  - When possible, QC SNPs within chips for better performance
- Included in ICAR guidelines on May 20th, 2026
  - *Section 04 – DNA Technology*,
    - *Chapter 6 – "Parentage Verification Using Full SNP Comparisons"*



Complementary approach to ISAG PV, to be preferred when available.

Acknowledgements

- **ICAR DNA** working group
- Participating **dairy producers** for supplying data
- **DHI organizations and DRPCs** for processing and relaying the information to CDCB
- **Purebred breed associations** for providing pedigree data

Thank you for your attention!

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 prover and employer.

Validation

- Independent(*) population HO/JE
- ~60k comparisons each
- 99.75% sensitivity, 99.95% specificity

DUO comparisons				NON DUO comparisons		
Threshold	Total	Parentage		Threshold	Total	Parentage
0.0%	12745	ACCEPTED		< 0.6%	0	ACCEPTED
0.01% ≥ 0.1%	47454	ACCEPTED		0.61% ≥ 0.7%	0	DOUBTFUL
0.11% ≥ 0.2%	1988	ACCEPTED		0.71% ≥ 0.8%	4	DOUBTFUL
0.21% ≥ 0.3%	280	ACCEPTED		0.81% ≥ 0.9%	8	DOUBTFUL
0.31% ≥ 0.4%	88	ACCEPTED		0.91% ≥ 1%	12	DOUBTFUL
0.41% ≥ 0.5%	39	ACCEPTED		1.01% > 2%	1373	EXCLUDED
0.51% ≥ 0.6%	25	ACCEPTED		2.01% > 3%	3324	EXCLUDED
0.61% > 1%	103	DOUBTFUL		> 3%	53582	EXCLUDED
> 1%	62	EXCLUDED				

- Most of >1% results likely linked to DNA quality or density
  - 40's/50's/60's bulls
  - embryo genotypes (method)
  - Few animals with low density chips (<10k SNPs)

When using QC SNPs **on not desired results**:
not accepted in DUO:
14 DOUBTFUL,
4 EXCLUDED (3 LLD chip, 1 embryo)
not excluded in NON DUO: 2 DOUBTFUL