



IRISH CATTLE BREEDING FEDERATION

Challenges of using SNPs for parentage confirmation and parentage determination in beef and dairy cattle:
Nationwide perspective when genotyping commercial and pedigree animals

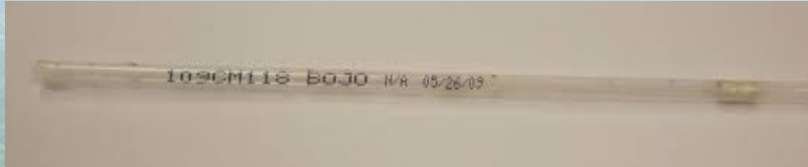
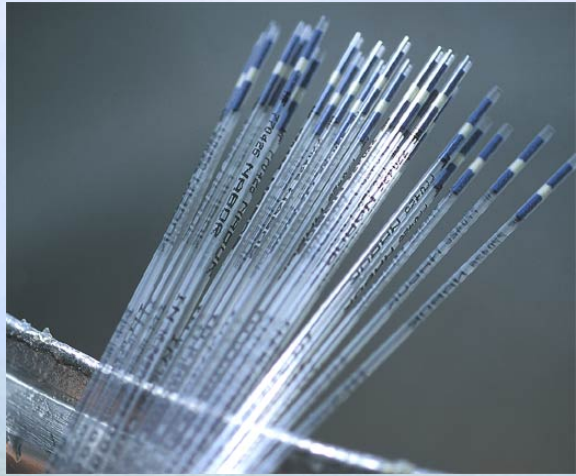


Matthew McClure, ICBF

Interbull , Verden, Germany 26/02/2015



Parentage Validation/Prediction



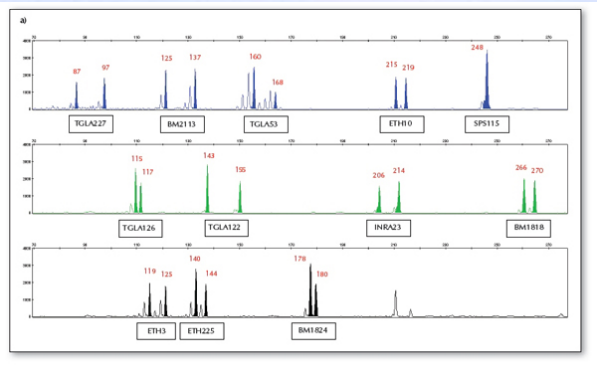
Parentage Validation/Prediction

How many needed?

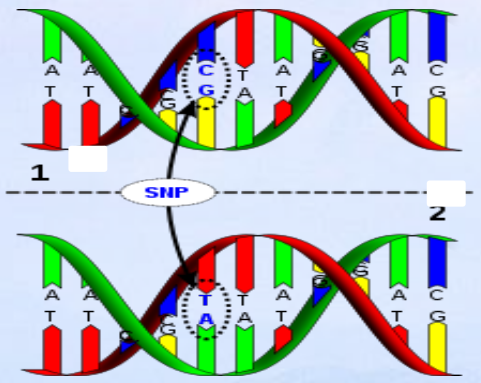
Recipient's blood			Reactions with donor's red blood cells			
ABO antigens	ABO antibodies	ABO blood type	Donor type O cells	Donor type A cells	Donor type B cells	Donor type AB cells
None	Anti-A Anti-B	O				
A	Anti-B	A				
B	Anti-A	B				
A & B	None	AB				

Compatible Not compatible

Blood Typing



Microsatellite Markers



SNPs

Microsatellites

ISAG Panel: BM1824, BM2113, INRA023, SPS115, TGLA122, TGLA126, TGLA227, ETH10, ETH225, ETH3, SPS115, TGLA53

Additional panel of MGTG4B, CSRM60, SPS113, ILSTS006, RM067, CSSM66

SNP: ISAG100
ISAG200
Some labs use more

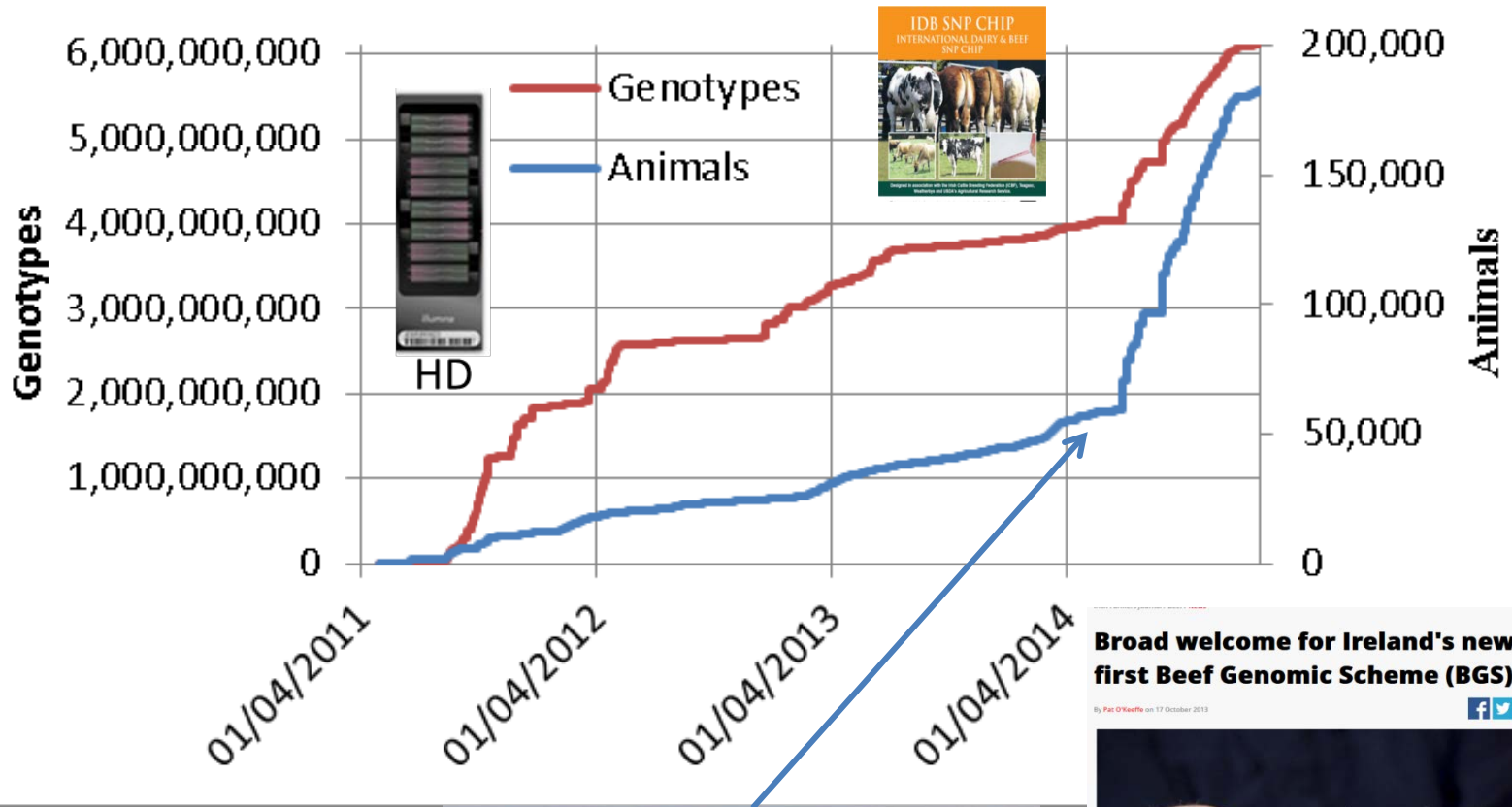


121 Validation
2000 Prediction

How Many SNP to Predict 1 Sire



ICBF Database Growth



Broad welcome for Ireland's new world first Beef Genomic Scheme (BGS)

By Paul O'Keefe on 17 October 2013

How Many SNP to Predict 1 Sire

Sire Validation Errors with Small SNP Sets

- Fail Sire A with 121 SNP (ISAG200 subset)
 - >1 misconcordance = fail
- Predict Sire A with 2,000 SNP
 - <10 misconcordance = pass



Unsort	Map	itt	sire_itt	ARS-USMARC-PARENT-EF093511-RS29011320
1	HERAUSM00000GRKA584		1	B_B
2	HERIRLM141607070601	HERAUSM00000GRKA584		A_A
3	HERIRLM151225330379	HERAUSM00000GRKA584		A_A
4	HERIRLM151225340413	HERAUSM00000GRKA584		A_A
5	HERIRLM151225360423	HERAUSM00000GRKA584		A_A
6	HERIRLM151225370399	HERAUSM00000GRKA584		A_A
7	HERIRLM151225380425	HERAUSM00000GRKA584		A_A
8	HERIRLM151225380441	HERAUSM00000GRKA584		A_A
9	HERIRLM191141760568	HERAUSM00000GRKA584		A_A
10	HERIRLM231305310207	HERAUSM00000GRKA584		A_A
11	HERIRLM241245190385	HERAUSM00000GRKA584		A_A
12	HERIRLM331493150493	HERAUSM00000GRKA584		A_A
13	HERIRLM361055830685	HERAUSM00000GRKA584		A_A
14	HERIRLM361055880665	HERAUSM00000GRKA584		A_A
15	HERIRLM121182590084	HERAUSM00000GRKA584		A_B
16	HERIRLM121201450915	HERAUSM00000GRKA584		A_B
17	HERIRLM121201460899	HERAUSM00000GRKA584		A_B
18	HERIRLM141643520372	HERAUSM00000GRKA584		A_B
19	HERIRLM151225310435	HERAUSM00000GRKA584		A_B
20	HERIRLM151225330412	HERAUSM00000GRKA584		A_B
21	HERIRLM151225390376	HERAUSM00000GRKA584		A_B
22	HERIRLM151313160329	HERAUSM00000GRKA584		A_B
23	HERIRLM151654470902	HERAUSM00000GRKA584		A_B
24	HERIRLM211463520049	HERAUSM00000GRKA584		A_B
25	HERIRLM211463580046	HERAUSM00000GRKA584		A_B
26	HERIRLM231305370204	HERAUSM00000GRKA584		A_B
27	HERIRLM241842660174	HERAUSM00000GRKA584		A_B
28	HERIRLM281110080573	HERAUSM00000GRKA584		A_B
29	HERIRLM291166240641	HERAUSM00000GRKA584		A_B
30	HERIRLM301207290630	HERAUSM00000GRKA584		A_B
31	HERIRLM311122080052	HERAUSM00000GRKA584		A_B
32	HERIRLM311317670240	HERAUSM00000GRKA584		A_B
33	HERIRLM331340371248	HERAUSM00000GRKA584		A_B
34	HERIRLM331358460991	HERAUSM00000GRKA584		A_B
35	HERIRLM121182560081	HERAUSM00000GRKA584		B_B
36	HERIRLM141607090603	HERAUSM00000GRKA584		B_B
37	HERIRLM151225340421	HERAUSM00000GRKA584		B_B
38	HERIRLM151225370424	HERAUSM00000GRKA584		B_B
39	HERIRLM151313130326	HERAUSM00000GRKA584		B_B



Sire Validation Errors with Small SNP Sets

- Fail with 121 then predict with 2,000 SNP
- Validate the wrong sire?



Could 121 SNP Validate Wrong Sire?

Animal	Listed sire		
LIMIRLF.....554	LIMIRLM.....050		
SNP set	#_mismatches	#SNP_both_genotyped	_%_SNP_mismatch
121 SNP	0	118	0.00%
LD_base	267	6848	3.90%
CHAIRLF.....639	CHAIRLM.....163		
SNP set	#_mismatches	#SNP_both_genotyped	_%_SNP_mismatch
121 SNP	1	113	0.88%
LD_base	336	6729	4.99%
CHAIRLF.....355	CHAIFRAM.....993		
SNP set	#_mismatches	#SNP_both_genotyped	_%_SNP_mismatch
121 SNP	1	114	0.88%
LD_base	227	6761	3.36%
LIMIRLF.....500	LIMIRLM.....959		
SNP set	#_mismatches	#SNP_both_genotyped	_%_SNP_mismatch
121 SNP	0	112	0.00%
LD_base	0	6711	0.00%

Sire Validation Errors with Small SNP Sets

- Fail with 121 then predict with 2,000 SNP
- Validate the wrong sire?
- **Why validate with small and predict large SNP**

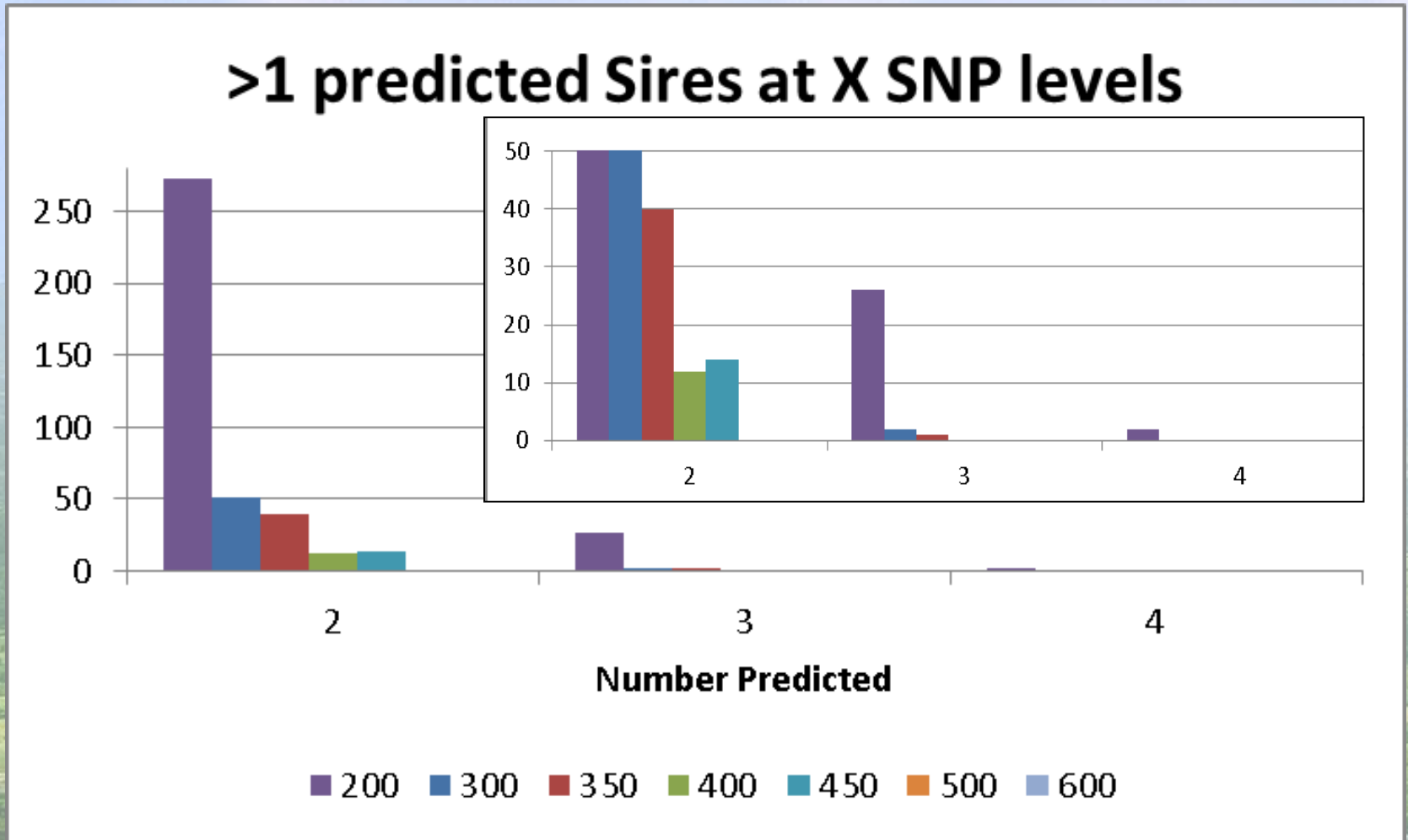


How Many SNP to Predict 1 Sire

- 56,000 genotyped animals
 - HD, 50K, IDB, LD
 - 3K genotypes excluded
- SNP densities tested
 - ISAG200
 - Plus top 100, 200,1800 MAF SNP
- 9,000 sire validated animals, 56K ref
- Predict sire/dams at 0.5-1% misconcordance rate

Animals	56,000
breed	Jun-14
HOL	68.65%
LIM	7.94%
CHA	7.41%
AAN	4.42%
SIM	2.35%
HER	3.06%
BBL	1.01%
MSH	0.06%
SAL	0.04%
JER	0.17%
PAR	0.17%
LMS	1.91%
BAQ	0.04%
AUB	0.00%
CHL	1.68%
PIE	0.55%
HFD	0.17%
MON	0.14%
IRM	0.00%
NWR	0.09%
RED	0.01%

Sire Prediction with 200, 300,... SNP



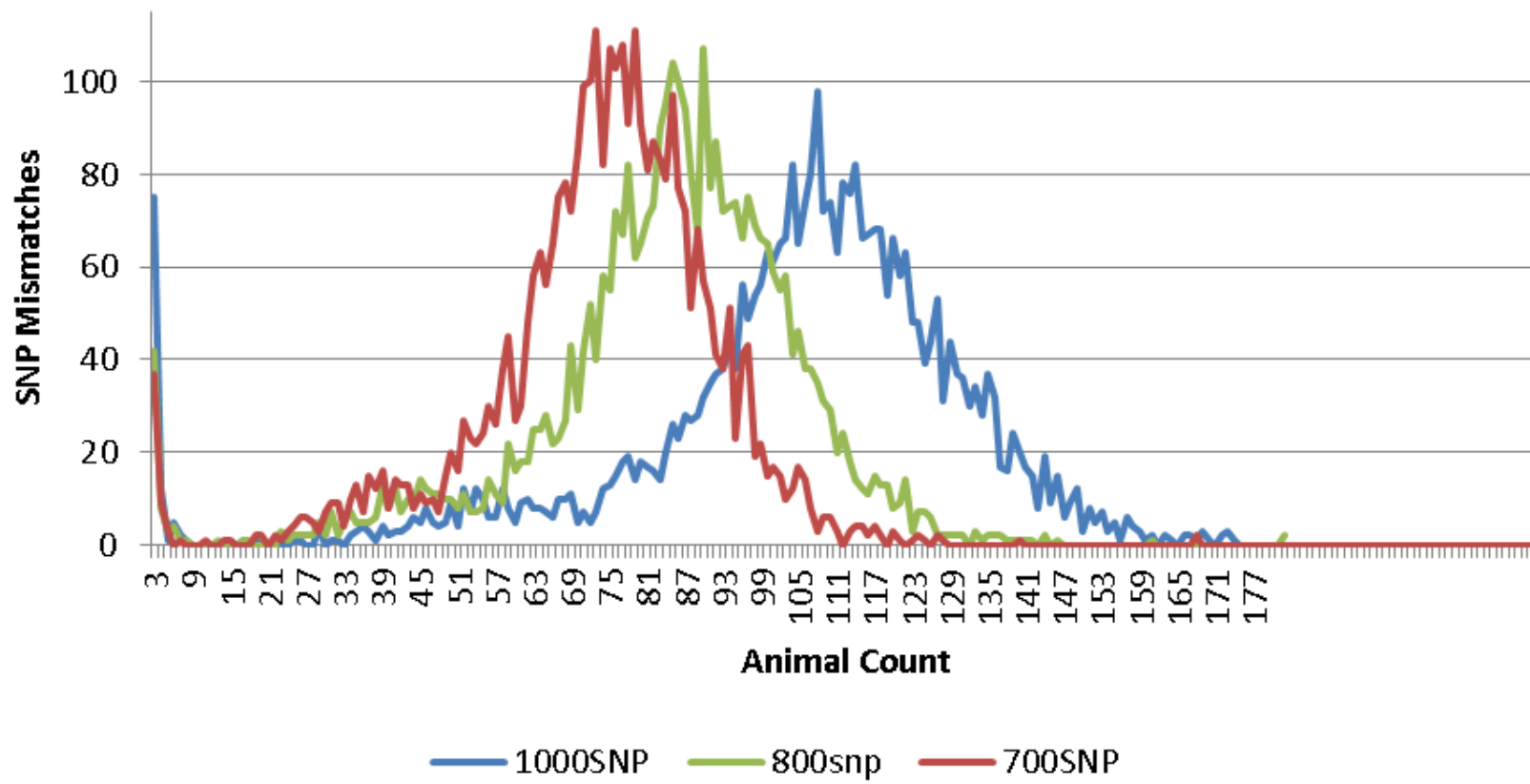
Space between Validation and Fail

2000 SNP

Validate

Fail

SNP Mismatches for sire validations 114,000 animals

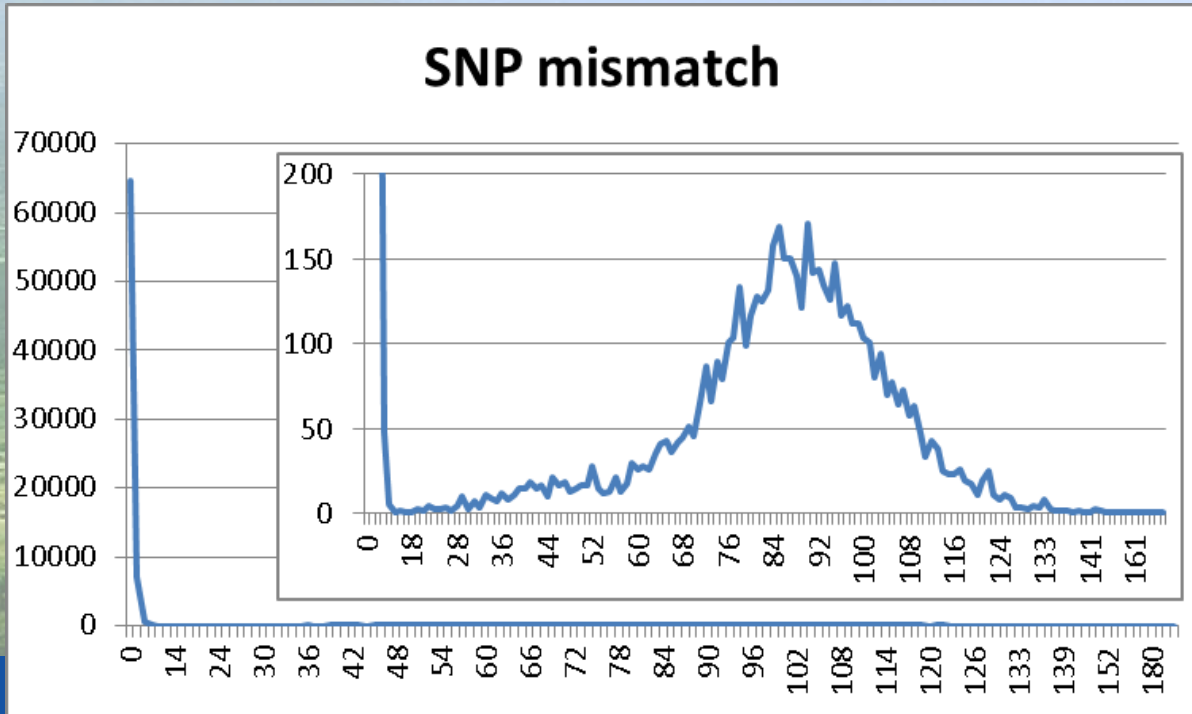


Could 121 SNP Validate Wrong Sire?

Animal	Listed sire		
LIMIRLF.....554	LIMIRLM.....050		
SNP set	# mismatches	#SNP both genotyped	% SNP mismatch
121 SNP	0	118	0.00%
800 SNP	38	800	4.75%
LD_base	267	6848	3.90%
CHAIRLF.....639	CHAIRLM.....163		
SNP set	# mismatches	#SNP both genotyped	% SNP mismatch
121 SNP	1	113	0.88%
800 SNP	38	777	4.89%
LD_base	336	6729	4.99%
CHAIRLF.....355	CHAFRAM.....993		
SNP set	# mismatches	#SNP both genotyped	% SNP mismatch
121 SNP	1	114	0.88%
800 SNP	18	782	2.30%
LD_base	227	6761	3.36%
LIMIRLF.....500	LIMIRLM.....959		
SNP set	# mismatches	#SNP both genotyped	% SNP mismatch
121 SNP	0	112	0.00%
800 SNP	0	781	0.00%
LD_base	0	6711	0.00%

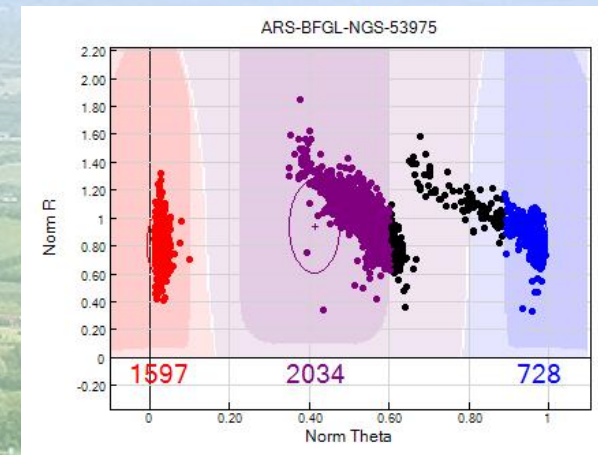
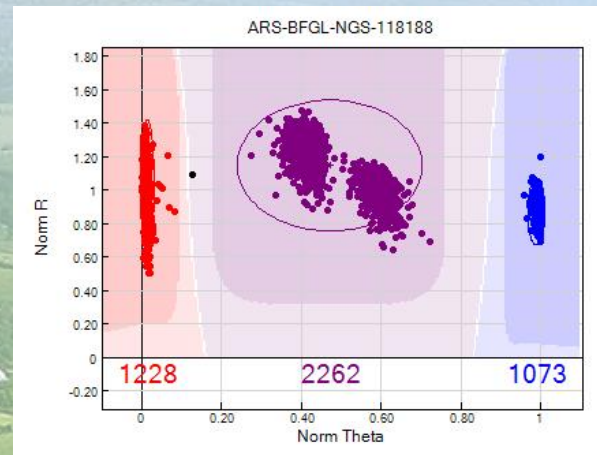
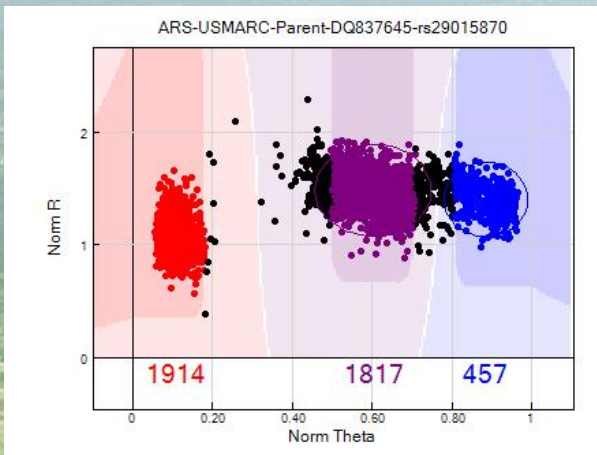
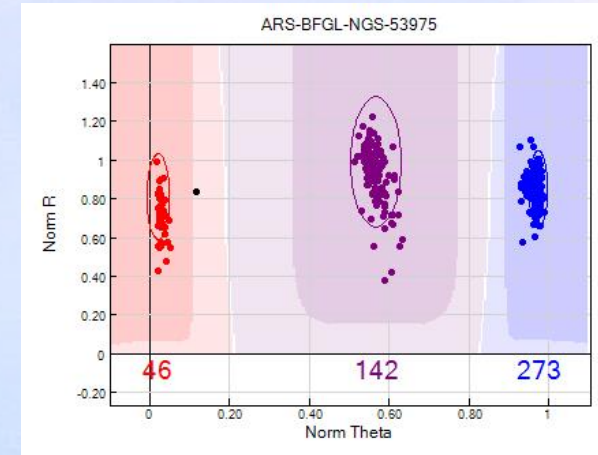
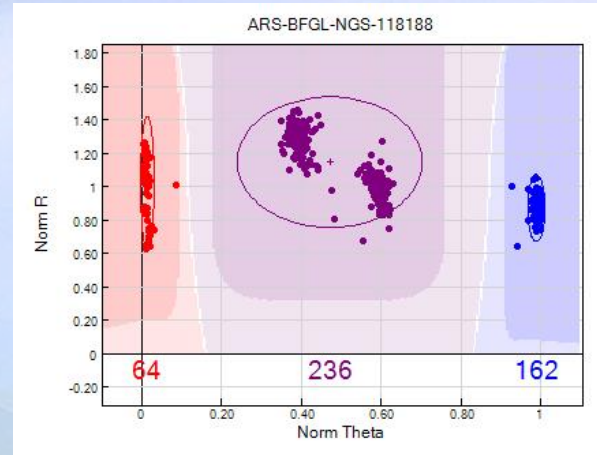
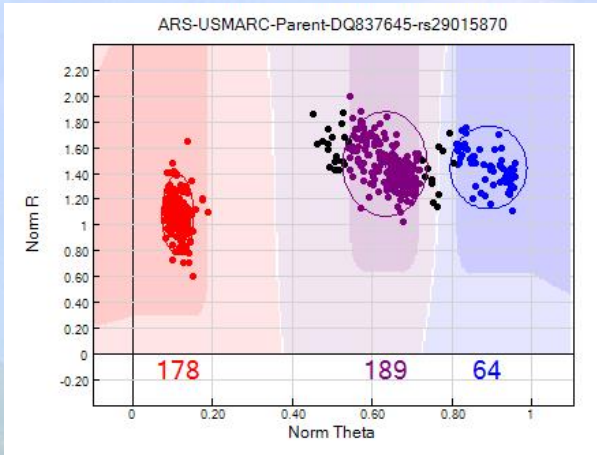
Use >200 SNP for Validation

- Switched to 800 SNP July 2014
- Refinement in February 2015



Animals	56,000	180,000
breed	Jun-14	Feb-15
HOL	68.65%	30.08%
LIM	7.94%	22.34%
CHA	7.41%	18.68%
AAN	4.42%	7.22%
SIM	2.35%	6.77%
HER	3.06%	4.66%
BBL	1.01%	2.75%
MSH	0.06%	1.57%
SAL	0.04%	0.92%
JER	0.17%	0.67%
PAR	0.17%	0.64%
LMS	1.91%	0.59%
BAQ	0.04%	0.59%
AUB	0.00%	0.53%
CHL	1.68%	0.52%
PIE	0.55%	0.17%
HFD	0.17%	0.07%
MON	0.14%	0.05%
IRM	0.00%	0.03%
NWR	0.09%	0.03%
RED	0.01%	0.03%

SNP Refinement QC Quality: Clusters when processing 400 vs 4K / week



ISAG 2014 Workshop Report

ISAG100/200 SNP with platform consensus <99%

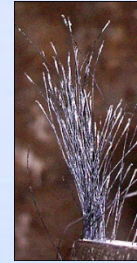
Marker Name	1 st	2 nd	Average
ARS-USMARC-Parent-DQ786764-no-rs	85.64	99.72	92.68
ARS-USMARC-Parent-DQ786766-rs29012070	96.08	90.22	93.15
ARS-USMARC-Parent-AY761135-rs29003723	92.59	97.22	94.91
ARS-USMARC-Parent-DQ846691-rs29019814	98.44	96.20	97.32
ARS-USMARC-Parent-EF034087-no-rs	95.10	100.00	97.55
ARS-USMARC-Parent-AY851163-rs17871661	100.00	95.24	97.62
ARS-USMARC-Parent-DQ990834-rs29013727	97.20	99.72	98.46
ARS-USMARC-Parent-EF034081-rs29009668	97.54	99.72	98.63
ARS-USMARC-Parent-DQ647187-rs29010510	97.56	99.72	98.64
ARS-USMARC-Parent-EF042090-no-rs	97.37	100.00	98.69
ARS-USMARC-Parent-DQ839235-rs29012691	100.00	97.50	98.75
ARS-USMARC-Parent-EF034082-rs29013532	97.74	100.00	98.87
ARS-USMARC-Parent-AY842474-rs29003226	98.15	99.72	98.94

1 reference and <20 samples

Genotyping platforms: KASPar, Illumina, Fluidigm, Sequenome, USDASNP96, Open Array, NGS

Animal Considerations

- Pedigree and Commercial animals
- Duplicate Genotypes
 - Few hundred out of 180,0000
- Sex off
 - Angry Bulls or registration error?
- Breed composition check



Angus

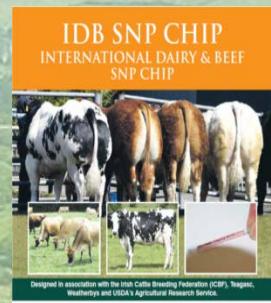
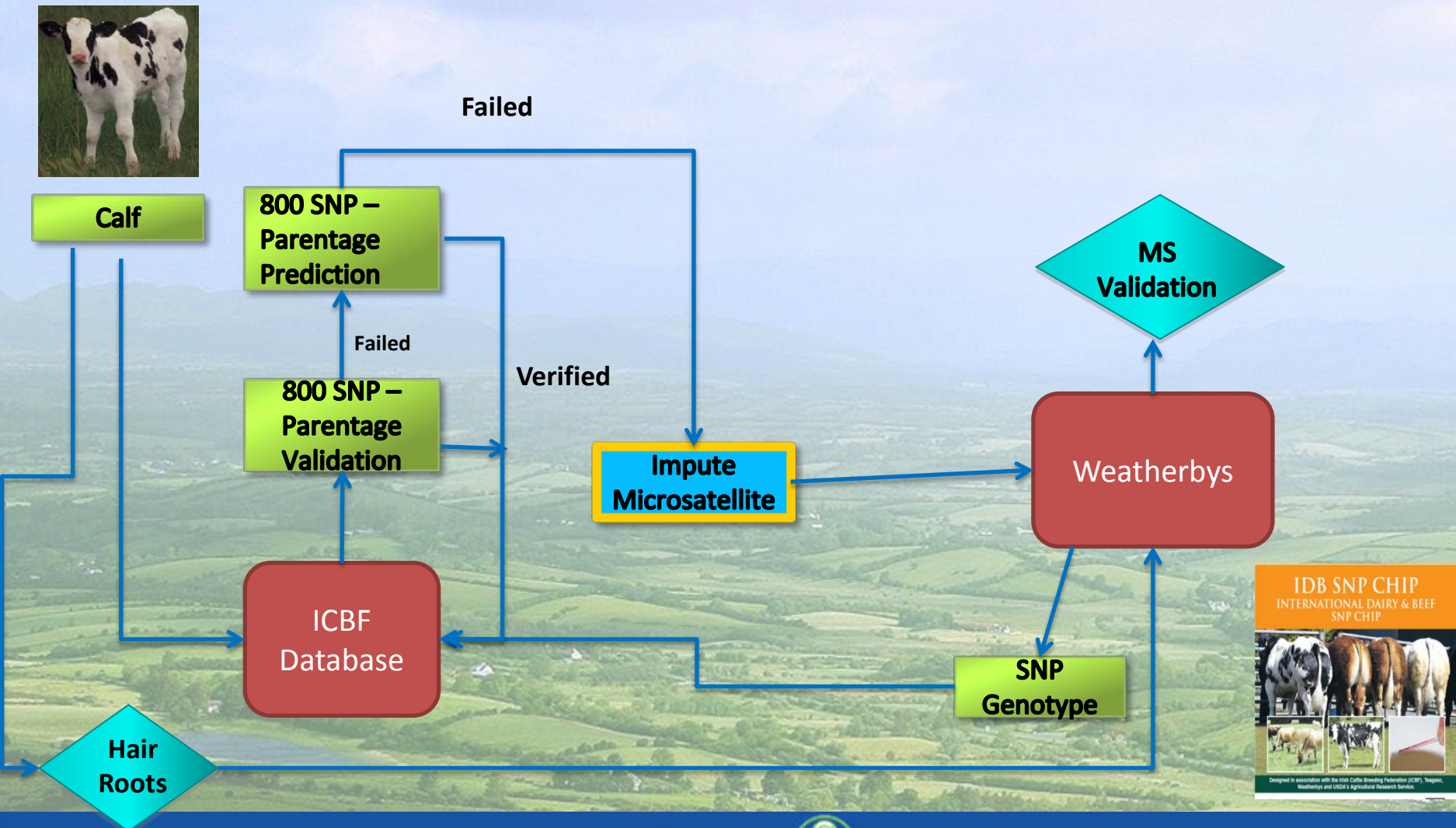


Limousin



Dexter

Irish Parent Verification Process



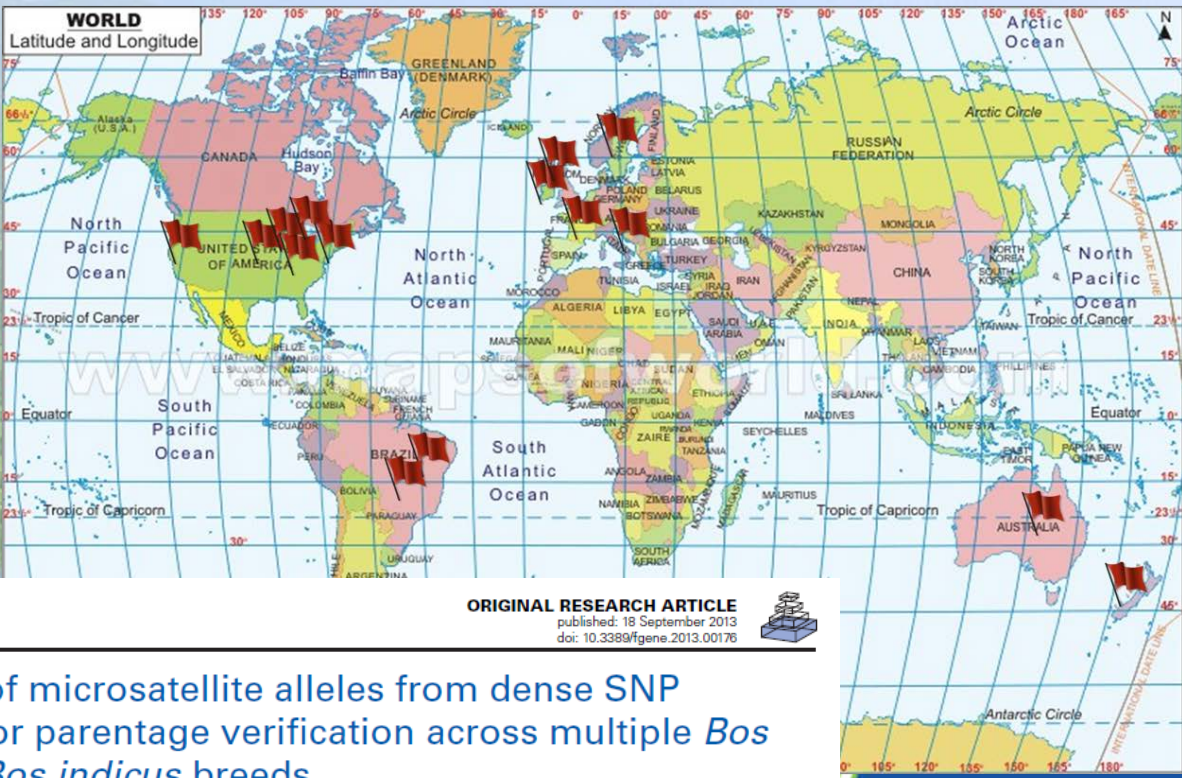
Microsatellite Imputation

985 SNP

Reference Population=>8,000

~2 hours to impute 4,000 in Beagle

>~16,000 MS imputed Irish animals (May, 2014)



frontiers in GENETICS

ORIGINAL RESEARCH ARTICLE
published: 18 September 2013
doi: 10.3389/fgene.2013.00178



Imputation of microsatellite alleles from dense SNP genotypes for parentage verification across multiple *Bos taurus* and *Bos indicus* breeds

Breed	SNP	SNP+MS
Abundance	172	165
Angus	623	235
Aubrac	239	234
Ayshire	523	86
Bazadaise	80	53
Beefmaster	36	36
Belmont Red	40	
Belgian Blue	210	12
Blonde D'Aquitain	225	201
Brahman	410	364
Brangus	13	13
Braunvieh	17	17
Bretonne Pie Noire	27	16
Brown Swiss	91	64
Brune Des Alpes	109	109
Charolais	1449	1109
Chiangus	19	0
Crossbred	506	3
Devon	16	16
Dexter	15	15
Fresian	163	35
Gasconne	142	142
Gelbvieh	44	0
Gir	209	101
Guernsey	110	18
Hereford	853	243
Holstein	2596	678
Jersey	87	131
Kerry	1	0
Limousin	2171	1572
Longhorn	13	13
Maine Anjou	38	16
Montbéliarde	257	251
Murray Grey	22	
N'Dama	24	0
Nelore	2659	135
Normande	256	242
Parthenaise	291	218
Pie Rouge Des Plaines	160	116
Piedmontese	24	17
Red Angus	61	47
Red Pie (italian)	2	0
Romagnola	23	23
Rouge Flamande	41	41
Salers	258	234
Santa Gertrudis	99	0
Sheko	18	0
Shorthorn	188	0
Simmental	907	324
Swedish Red	12	9
Tarentaise	167	155
Tropical Composite		336
Vosgienne	53	49
Watusi	15	15
TOTAL	17138	8259

Commercial herds

BGS Males	N=21,000
Sire SNP validate	55.00%
SNP Predict	4.70%
SNP Fail	2.50%
MS Validate	16.50%
MS Fail	1.50%
Sire not Genotype	20.00%
Overall Males	
Sire Validate	76%
Sire Fail	4%
Limbo	20%

Operational Considerations

- Building a parentage validation pipeline
 - including MS imputation
- Parentage prediction
 - Computation time and reference size
- Cost of STR's versus SNP's
 - Tipping point on cost and use (prediction/sex/carriers)
- Dealing with 'old' pedigree errors

Thank You

Chr	800	200	MS
1	49	11	62
2	47	9	55
3	35	8	97
4	27	9	0
5	39	9	73
6	23	8	0
7	34	11	0
8	35	6	0
9	22	6	104
10	31	9	0
11	32	10	0
12	23	5	0
13	32	6	0
14	24	5	0
15	18	7	54
16	25	5	125
17	31	9	0
18	28	7	54
19	19	8	128
20	16	5	89
21	31	8	77
22	25	7	0
23	18	4	67
24	26	4	0
25	20	4	0
26	18	4	0
27	19	6	0
28	21	6	0
29	18	4	0

