Genomic predictions - Current routines and experiences

Gert Pedersen Aamand, Ulrik Sander Nielsen
and Jørn Pedersen

Genotyped bulls

Birth year	HOL		RDC		Jersey	
	HB	No HB	HB	No HB	HB	No HB
2007	356	94	219	0	55	4
2008	313	524	257	30	47	33
2009	210	835	246	333	58	130
2010	207	951	213	787	40	210
2011	9	(1040)	95	(945)	0	(256)
Total all years	**6756**	**3645**	**5764**	**2096**	**41**	**623**

Selection intensity increases by birth year

Genotyped females

Birth year	HOL	RDC	Jersey
<2007	524	131	105
2007	259	122	73
2008	461	223	1040
2009	623	387	135
2010	953	629	122
2011	361	326	22
Total	3181	1818	561

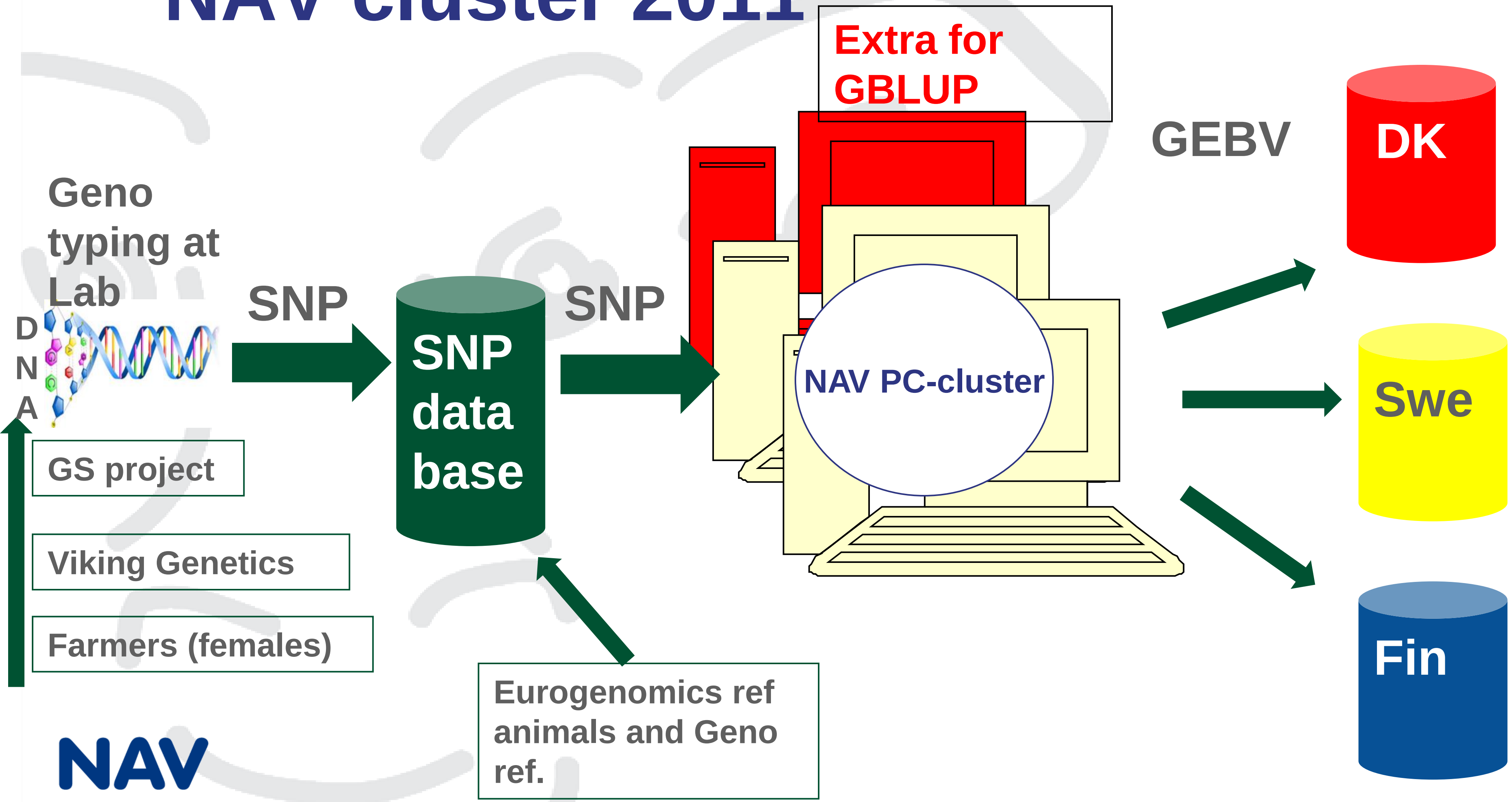
Genotyped females, RDC

Birth year	DNK	FIN	SWE
2007	30	68	24
2008	58	135	30
2009	78	252	57
2010	145	394	90
2011	41	253	32
Total	385	1159	274

Genotyped females, HOL

Birth year	DNK	FIN	SWE
2007	175	32	50
2008	338	40	74
2009	395	99	122
2010	638	224	91
2011	212	112	37
Total	2142	531	466

NAV cluster 2011



Geno typing at Lab

DNA

GS project

Viking Genetics

Farmers (females)

SNP

SNP data base

Eurogenomics ref animals and Geno ref.

SNP

Extra for GBLUP

NAV PC-cluster

GEBV

DK

Swe

Fin

GEBV

Remember

- A new method - start 2008
- Development very fast
- Very fast movement from R&D to implementation in practice
- A lot of changes for NAV, national databases, breeding organizations and farmers

GEBV

Since 2008/09	Unofficial genomic breeding values
2nd May 2011	First official GEBVs for NTM traits
June 2011	First official intermediate run – GEBV females updated
9th August 2011	First official GEBVs for all type traits
2nd Nov 2011	First official GEBVs for DRH

NAV routine evaluations

Normal routine runs

2nd May, 9 August, 2 November, 2 February

- **EBVs based on traditional information**
- **EBVs combining traditional information+genomic information**

Extra routine runs with GEBV

15 June, 15 September, 15 December ,15 March

- **EBVs combining traditional information+genomic information**

Official GEBVs November 2011

Traits

- NTM traits + linear type traits

Category of <u>genotyped</u> animals

- AI bulls (no progeny test yet)
- Heifers
- Cows (traits based on pedigree only)

Breeding values!

November 2011 we have:

- EBVs – based on phenotype (lots of animals)
- EBVs – based on phenotype and genomic information (a few animals)

The EBVs from the two "categories" are comparable and the SD reflects the reliability

Publication of breeding values

Category of animals	Status	Published breeding value
Bulls without a progeny test	Culled	None
	AI bulls with a Nordic herd book number	**GEBV** when at least 20 month old at publication date
Bulls with Nordic or foreign progeny test	AI bulls with a Nordic progeny test	EBV
	Foreign AI bulls with a Nordic herd book number and a progeny test abroad	IB EBV for all international traits available. **GEBV** for traits with pedigree information only

Publication of GEBV (genotyped animals)

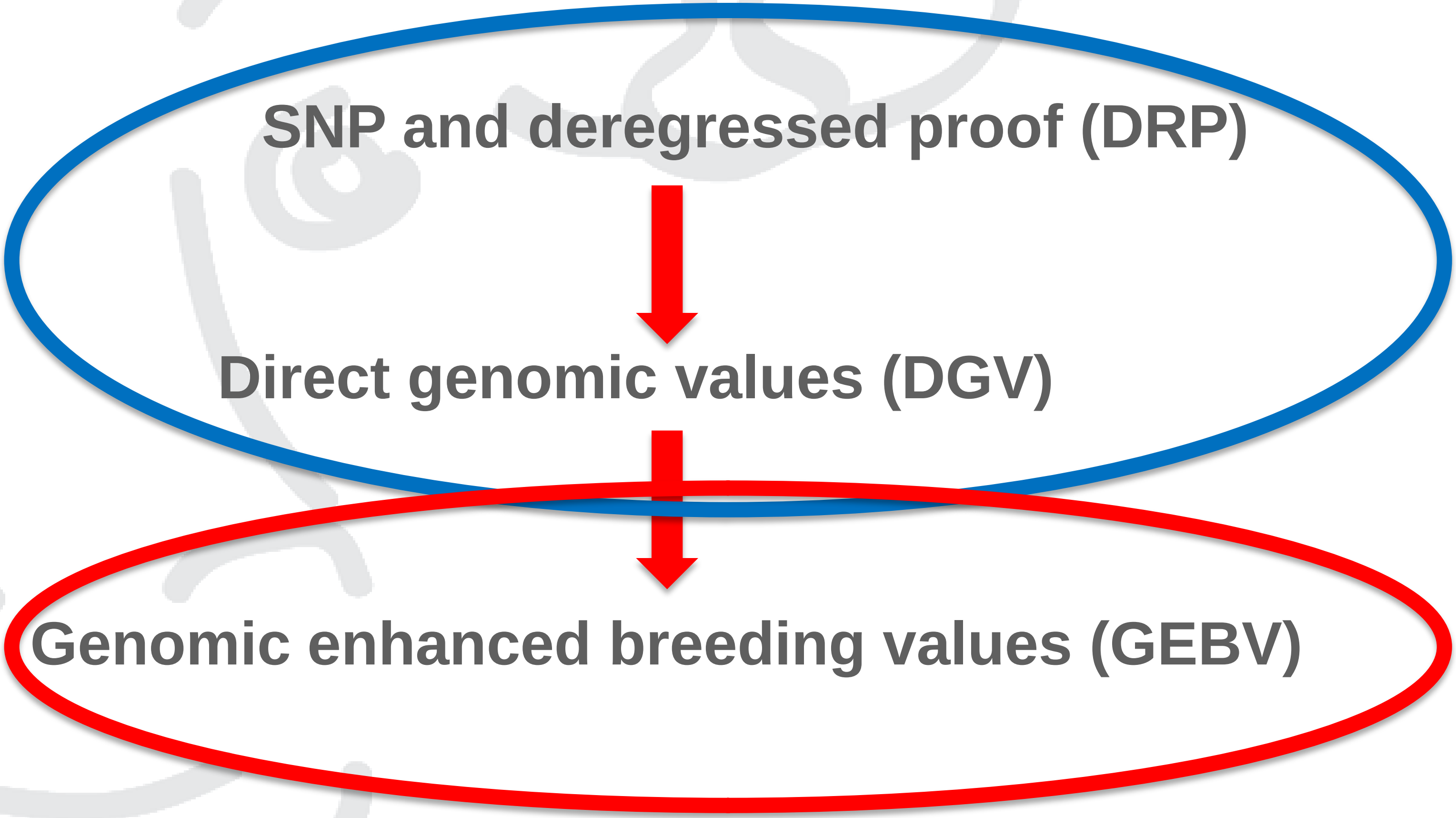
Category of animals	Published Breeding value
Heifers	**GEBV**
Cows	**GEBV** for traits with pedigree information only (e.g. other disease, fertility, calving) and EBVs for all other traits

Method

NAV uses a so-called 2-step method:

- GBLUP for estimation of DGVs
- A two trait model to combine DGVs and phenotypic information (DRP) (Mäntysaari, 2010)

Current 2-step method



SNP and deregressed proof (DRP)

Direct genomic values (DGV)

Genomic enhanced breeding values (GEBV)

Input data: Deregressed proofs

Data from practice (kg protein, mastitis diagnoses)

↓ **Systematic effects**
(calving age, calving month etc.)

EBV

↓ **Reliability taken into account**

Deregressed proof (DRP)

Deregressed proof (DRP) basis for DGV

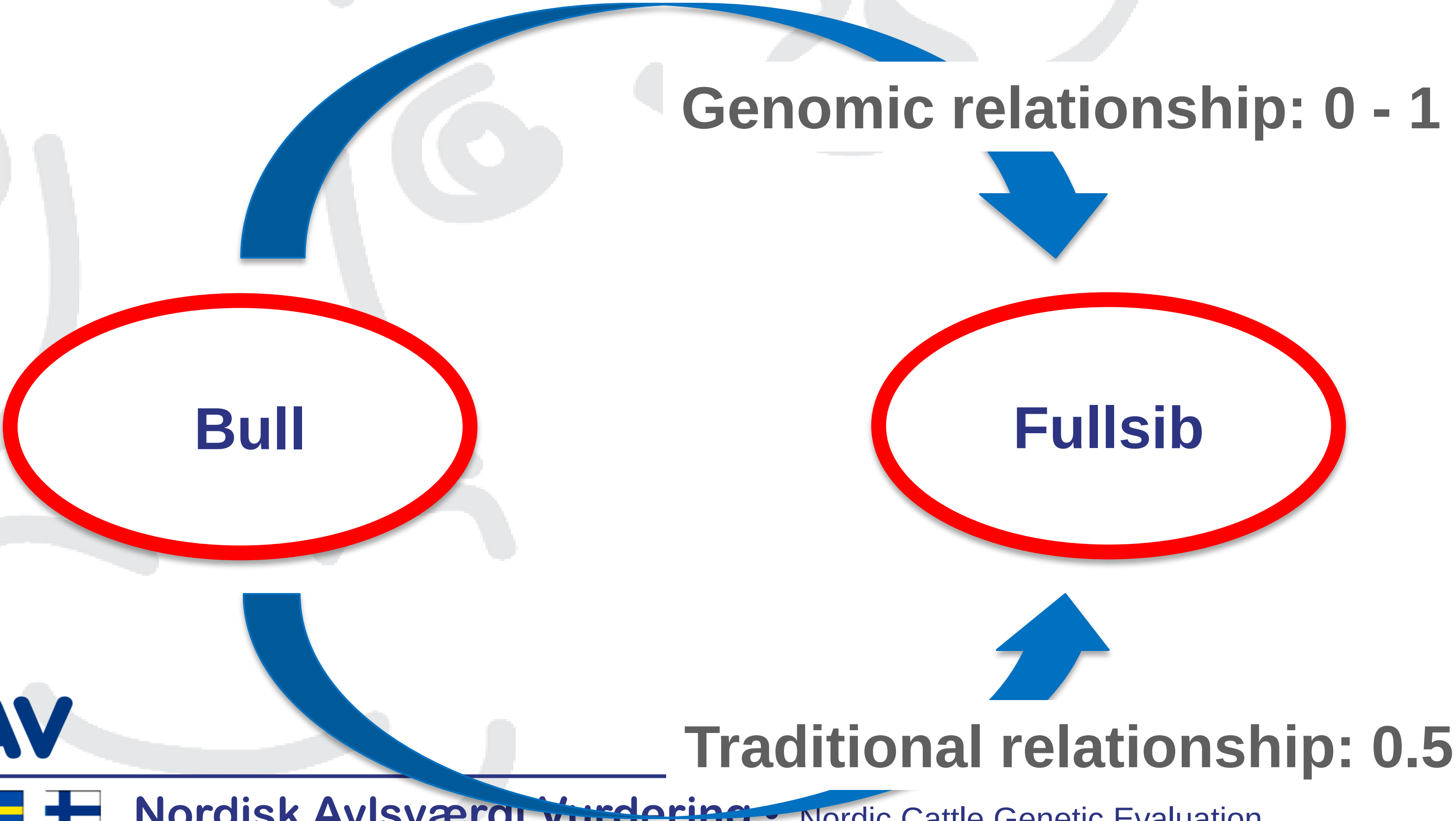
Deregressed proof
or phenotypes

Estimation of BV

Same
EBVs

DRP combines a lot of phenotype data in one figure per bull

Relationship based on SNP



Genomic relationship: 0 - 1

Bull

Fullsib

Traditional relationship: 0.5

Genomic prediction

DRP = μ + genotyped animal + e

- Genomic relationship
- Bulls without DRP included
- DRP reliability taking into account in the model
- No polygene effect

Resultat : DGV

Traditional estimation of breeding values

Phenotypic records = μ + fixed. effects + animal + e

- Traditional relationship matrix
- All females and bulls included
- Software and methods for genomic and traditional estimation of breeding values are basically the same!

Reliabilities of DGV for candidates are estimated from validation test

- Estimate DRP$_{-4years}$ including all genotyped bulls with progeny in the evaluation – the four youngest birth years are deleted afterwards
- Estimate DGV$_{-4år}$ based on DRP$_{-4year}$
- Keep DRP including all bulls (DRP$_{all}$)
- Do linear regression, including only bulls excluded from DRP$_{-4år}$
- Compare with a linear regression, substituting DGV with pedigree index

Lineær regression

Model 1:

$DRP_{all} = b_1 + b_2 * DGV_{-4year}$ (model 1)

- Expectation: $b_1 = 0$ og $b_2 = 1$
- Reliability DGV_{-4year} : R^2/ reliability DRP (R^2 model 1)

Model 2:

$DRP_{all} = b_1 + b_2(½\ Sire + ¼\ Mgs)$ (model 2)

- Reliability explained by pedigree: R^2/ reliability DRP (R^2 model 2)
- Proportion explained by pedigree:

Additional reliability = R^2(model 1) -R^2(model 2)

Reliability selected traits for HOL

	Rel. DRP	b_2	R^2	b_2	R^2	Rel. DGV	Rel. pedigree	Increase rel.
		Model 1		Model 2				
Protein	0.91	0.89	0.50			0.55		

Reliability selected traits for HOL

	Rel. DRP	b_2	R^2	b_2	R^2	Rel. DGV	Rel. pedigree	Increase rel.
		Model 1		Model 2				
Protein	0.91	0.89	0.50	1.04	0.23	0.55	0.25	0.30

Reliability selected traits for HOL

	Rel. DRP	b_2	R^2	b_2	R^2	Rel. DGV	Rel. pedigree	Increase rel.
		Model 1		Model 2				
Protein	0.91	0.89	0.50	1.04	0.23	0.55	0.25	0.30
Mastitis.	0.81	0.84	0.39	0.87	0.14	0,48	0.17	0.31
Legs	0.58	0.66	0.14	0.83	0.06	0.24	0,10	0.14

Reliability selected traits for RDC

	Rel. DRP	b_2	R^2	b_2	R^2	Rel. DGV	Rel. ped.	Increase rel.
		Model 1		Model 2				
Protein	0.92	0.82	0.30	0.83	0.10	0.33	0.11	0.22
Mast.	0.84	0.82	0.20	0.83	0.08	0.24	0.10	0.14
Legs	0.55	1.11	0.15	0.80	0.04	0.27	0.07	0.20

Reliability selected traits for JER

	Rel. DRP	b_2	R^2	b_2	R^2	Rel. DGV	Rel. ped.	Increase rel.
		Model 1		Model 2				
Protein	0.92	0.82	0.29		0.32	0.32	0,35	-0.03
Mast.	0.86	1.01	0.27		0.31	0.31	0.11	0.19
Legs	0.58	1.14	0.18		0.09	0.31	0.15	0.16

Impact on the size of reliability

- Number of animals in reference group
- Harmonization of traits – e.g. Euro Genomics, legs
- Homogeneity of reference group – e.g. Holstein versus RDC
- Reliability traditional EBVs
  - Heritability
  - Progeny group size

DGV information expressed as number of daughters

- Pedigree information included

	Reliability DGV									
Heritability	0.10	0.15	0.20	0.25	0.30	0,35	0.40	0.45	0.50	0.60
0.02	22	35	50	66	85	107	132	162	199	299
0.05	9	15	20	26	34	43	53	65	79	199
0.10	4	7	10	13	17	21	26	32	39	59
0.15	3	5	6	9	11	14	17	21	26	39
0.20	2	3	5	6	8	10	13	16	19	29
0.30	1	2	3	4	5	6	8	10	12	19

Pedigree and genomic information expressed as efficient daughters

example mastitis

Bull	HOL	RDC	Jersey
Total reliability DGV	38	26	34
Pedigeree	14	9	12
Extra information from GS	about 25	about 17	about 24

Relationship between reliability and standard deviation of EBV

Reliability = VAR(EBV)/VAR(genetic)

Eg. Yield HOL: 0.91 = $(10)^2/(\text{genetic std})^2$

Genetic std = 10.48

Std yield EBV with a given reliability has to be : √reliability * 10.48

STD DGV yield for HOL

- Std. on yield DGV from model HOL: 8.73
- DGV reliability from validation study: 0.55
- Correct std DGV: √0.55 *10.48 = 7.77
- DGV has to be corrected: 7.77/8.73 = 0.89
- Well known that DGV from model are inflated

STEP 2 Blending

Bivariate BLUP model: DRP og DGV (Mäntysaari 2010)

- DGV
  - Weight given to DGV: Reliability of DGV
- DRP
  - Weight given to DRP: Reliability DRP/EBV
- Relationship: From Sire and MGS (not AM)

Which animals are included in step II

Included animals:

- "Pedigree animals": Parents to animals with data (DRP and DGV)
- Reference bulls/progeny tested: Bulls with both DRP and DGV
- **Candidate bulls: Bulls with DGV information only (genotyped only)**
- **Heifer and cows: DGV only (genotyped)**

Published

- **Candidate bulls GEBV (> 20 mth)**
- **Heifer GEBV + Cow GEBV, except yield, mastitis and type**

Standardization of average

Average of GEBV = Average of EBV for:

- **Nordic AI bulls born 2001-2005**

Aprox. standard deviation on NTM

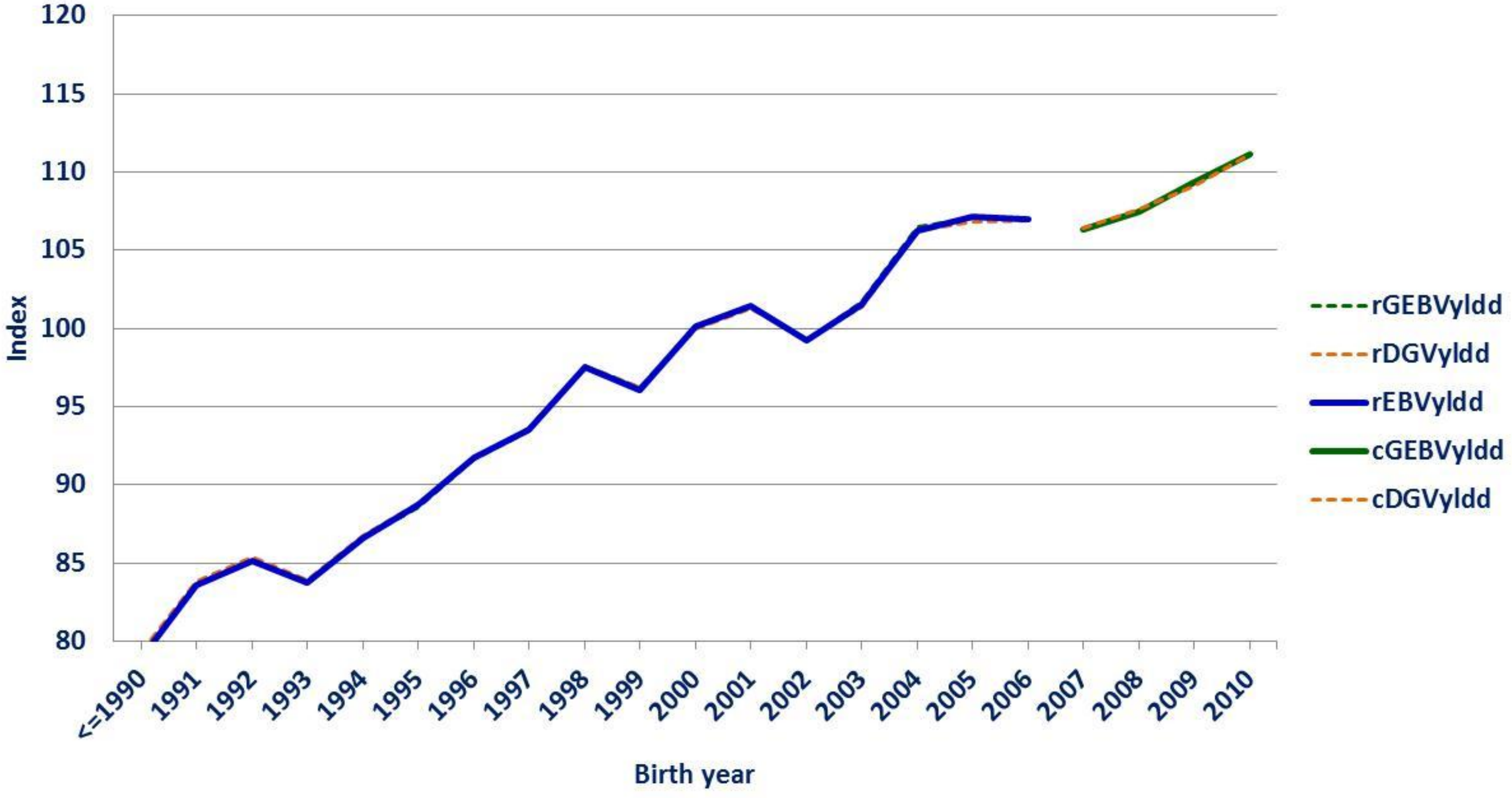
	Holstein	RDC	Jersey
Young bulls only genomic information	7-8	6-7	5-6
Progeny tested bulls	10	10	10

Higher SD corresponds to higher reliability
Reliability on genomic information across breeds:
Holstein> RDC > Jersey

HOL: GEBV, EBV and DGV trends for YIELD

Sires with genomic test, born in NAV-countries or AI-sires

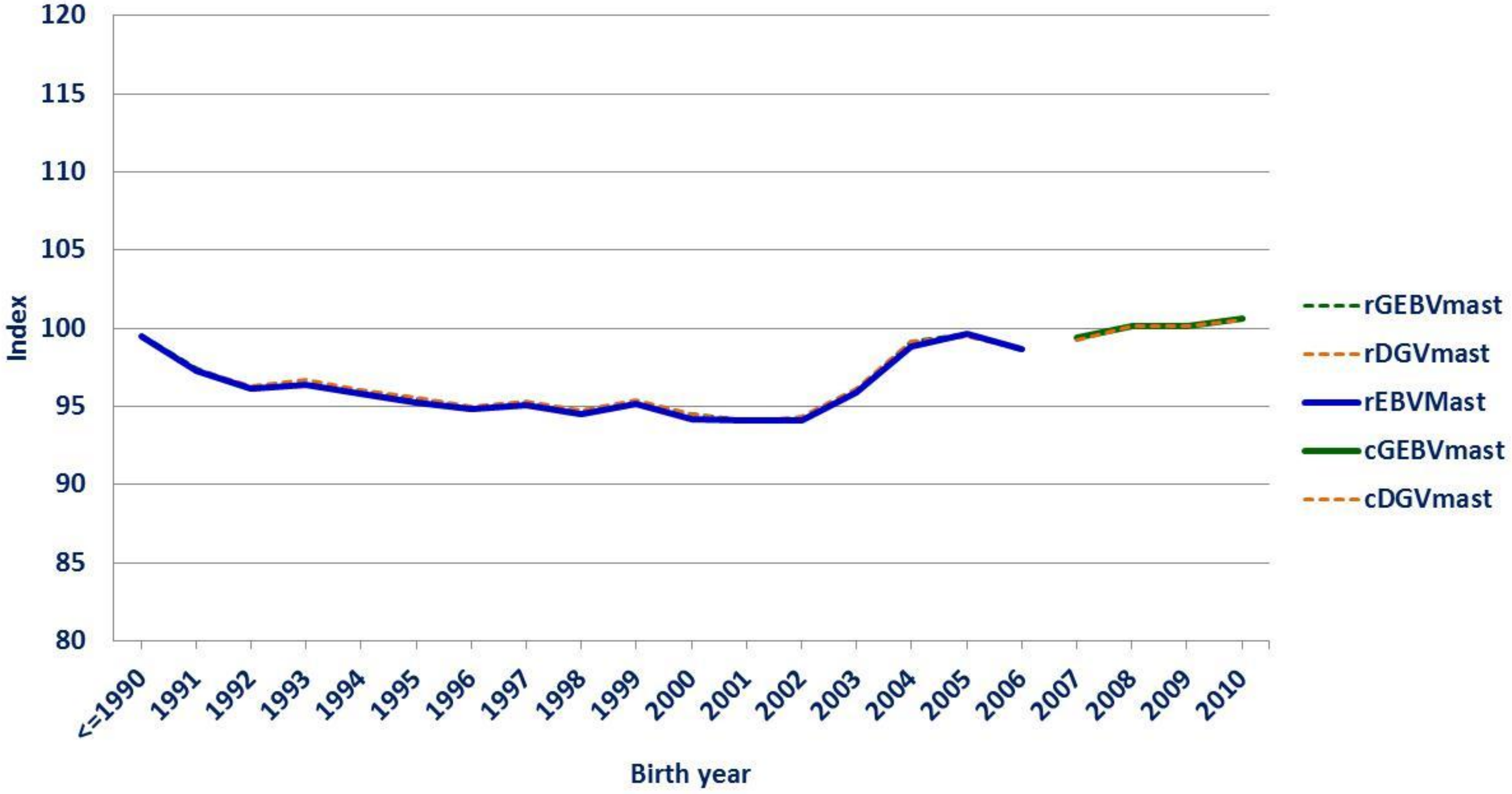
r = reference sire, c = candidate sire



HOL: GEBV, EBV and DGV trends for MASTITIS

Sires with genomic test, born in NAV-countries or AI-sires

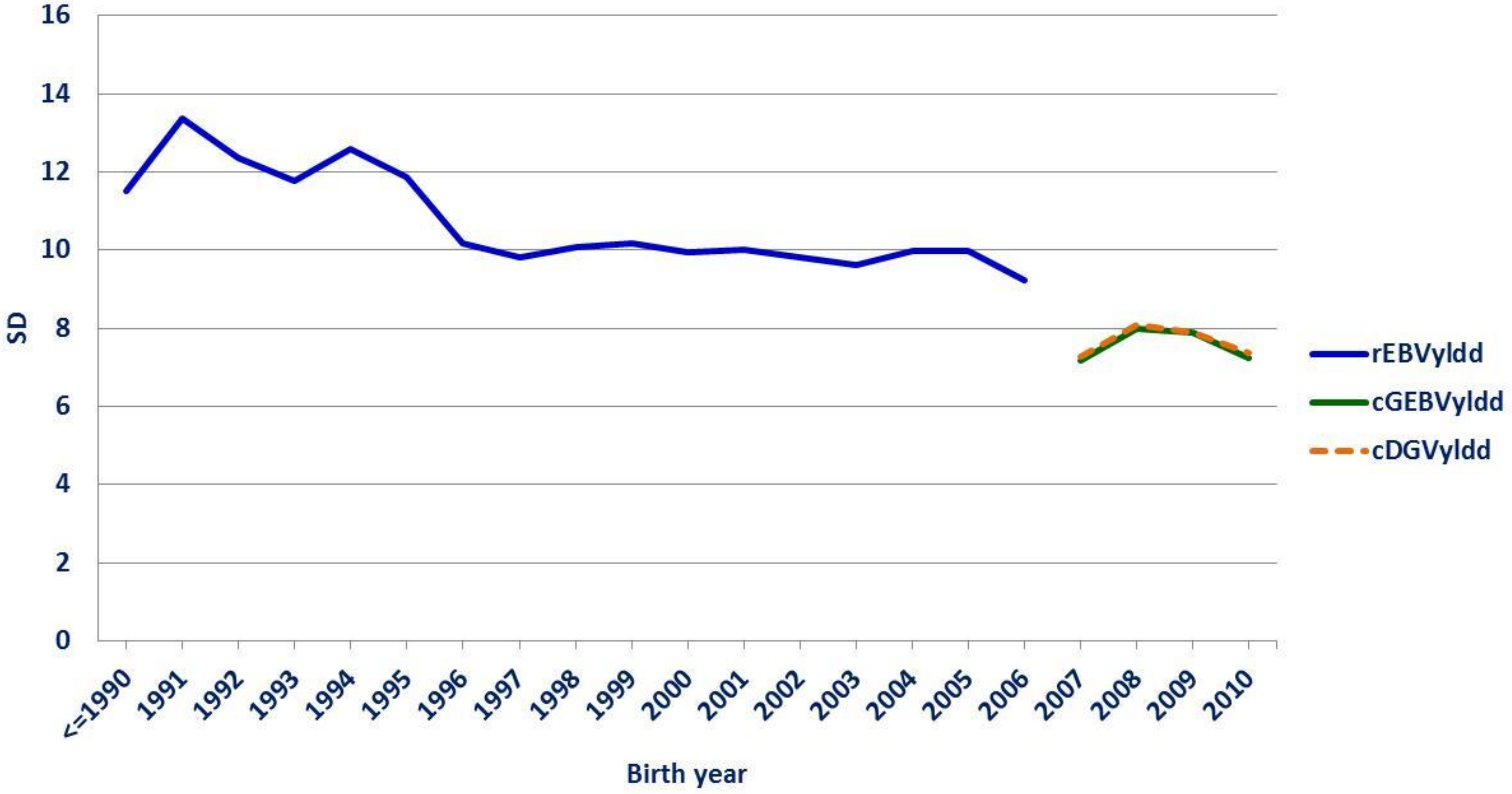
r = reference sire, c = candidate sire



HOL: Trend in SD of GEBV, EBV and DGV for YIELD

Sires with genomic test, born in NAV-countries or AI-sires

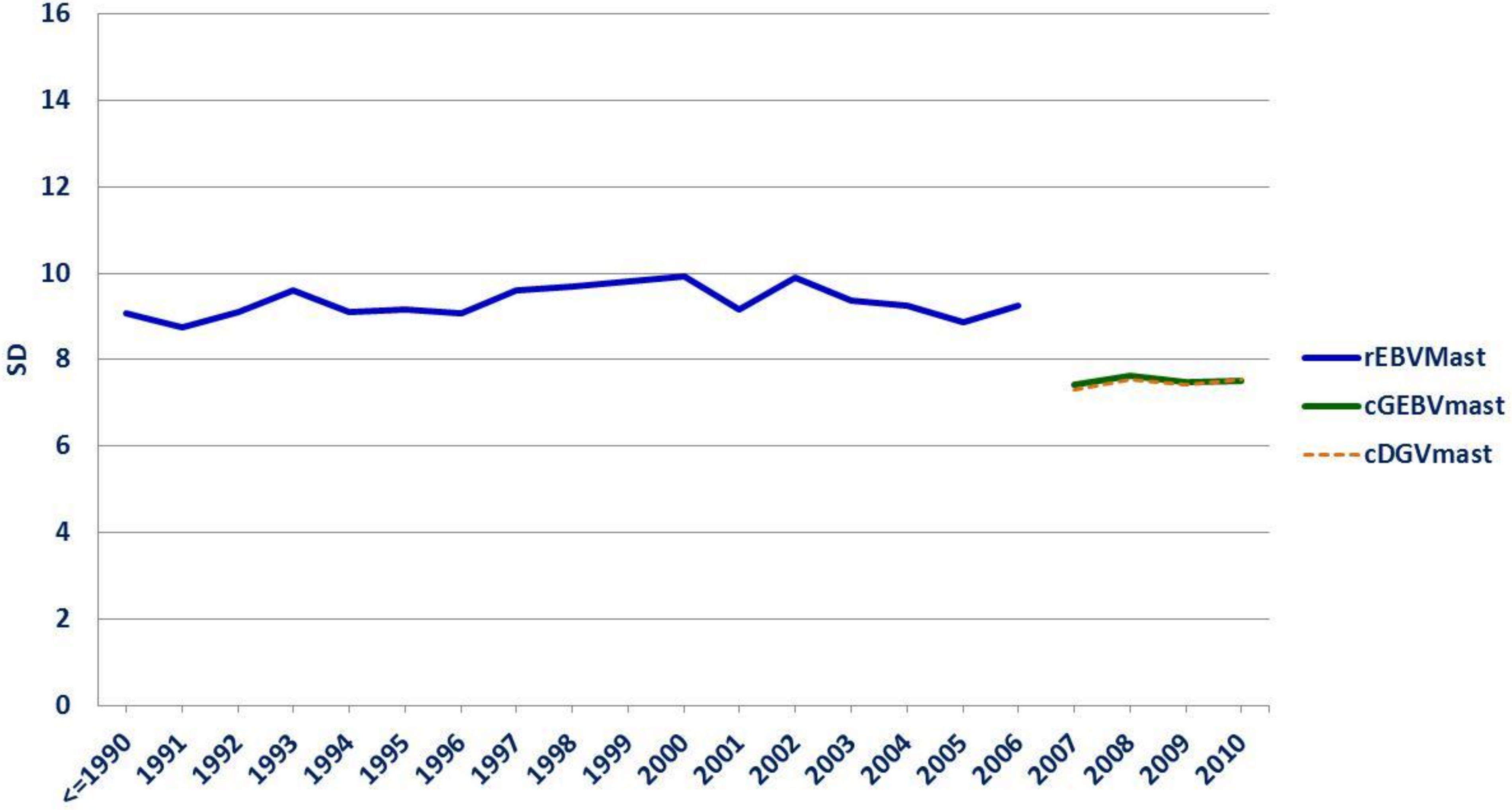
r = reference sire, c = candidate sire



HOL: Trend in SD of GEBV, EBV and DGV for MASTITIS

Sires with genomic test, born in NAV-countries or AI-sires

r = reference sire, c = candidate sire



Standard deviation of GEBV

Candidate sires

	Yield			Mastitis		
Birth year	HOL	RDC	JER	HOL	RDC	JER
2007	7.2	5.5	5.2	7.2	6.0	4.5
2008	7.8	5.9	5.6	7.6	5.8	6.7
2009	7.8	5.3	5.7	7.5	5.4	5.0
2010	7.8	5.2	5.8	7.7	5.7	5.0
Avg.	**7.7**	**5.5**	**5.6**	**7.5**	**5.7**	**5.3**

Correlations between GEBVs from August and November evaluation

Candidate sires in both evaluations (born 2007-2010)

	Yield	Fertility	Mastitis	Udder	Legs
HOL	0.994	0.988	0.989	0.997	0.993
RDC	0.990	0.985	0.984	0.996	0.994
JER	0.990	0.987	0.988	0.987	0.986

Candidate sire in previous and new evaluation - HOL
Distribution of GEBV changes for Yield

Birth year	Number	<-9	-9 to -8	-7 to -5	-5 to -4	-3 to-2	-1 to 1	2 to 3	4 to 5	6 to 7	8 to 9	>9	Avg Change
2007	377	0	0	0	0	55	313	9	0	0	0	0	-0.32
2008	837	0	0	0	0	175	634	28	0	0	0	0	-0.42
2009	1145	0	0	0	50	265	810	20	0	0	0	0	-0.66
2010	1156	0	0	0	10	182	937	27	0	0	0	0	-0.39
2011	774	0	0	0	1	107	643	22	1	0	0	0	-0.32

What happen, when a bull get (more) daughters?

- **Genomic tested bull**
  - **Reliability increase**
  - **Risk/chance for some changes**

- **Progeny tested bulls: More daughters**
  - **Reliability increase a little**
  - **Risk/chance for big changes is low**

- **Not genotyped bulls: Pedigree index substituted by pedigree index**
  - **Reliability increase a lot**
  - **Risk/chance for big changes**

Expected changes transition from young genomic tested bull to a progeny tested bull (95%)

Young bull

Protein: 100
Fertility: 100
NTM: 0

Progeny tested bull

Protein: 86 til 114
Fertility: 88 til 112
NTM: -13 til 13

Correlations between GEBV i Aug. - and EBV i Nov.

Candidate sires in AUG evaluation – reference sires in Nov. evaluation

Sires born 2006 and 2007

	Yield	Fertility	Mastitis	Udder	Legs
HOL (80 sires)	0.72	0.57	0.48	0.72	0.58
RDC (30 sires)	0.49	0.60	0.61	0.74	0.72
JER (10 sires)	0.46	0.71	0.82	0.62	0.39

Holstein bull birth year 2005-2007 Changes from candidate to reference bulls for yield

NOV	AUG/SEP	MAY	Number	EBV_{NOV}-$GEBV_{AUG}$	EBV_{NOV}-$GEBV_{MAY}$	$GEBV_{MAY-}GEBV_{AUG}$
REF	CAND	CAND	80	0.72	0.69	.
REF	REF	CAND	96	0.96	0.77	0.72

Candidate sire in previous - reference sire in new evaluation - HOL Distribution of GEBV changes for Yield

Birth year	Number	<-9	-9 to -8	-7 to -5	-5 to -4	-3 to-2	-1 to 1	2 to 3	4 to 5	6 to 7	8 to 9	>9	Avg Change
2006	20	1	1	0	3	2	1	3	4	2	1	2	1.57
2007	58	1	4	0	6	8	7	9	6	8	6	3	1.62

Genomic tested bull versus progeny tested – Holstein bulls

NAME	NTM	True NTM $_{95}$	
D Oscar	+32	+22 to + 42	160 daughters
VH OP	+32	+16 to +48	Genomic test

- **Note EBV has 50% chance to increase or decrease for both bulls**
- **Note the decrease or the increase can be larger when the reliability is lower – we are not able to estimate the EBV so precise**

Genotyped bulls

Birth year	HOL		RDC		Jersey	
	HB	No HB	HB	No HB	HB	No HB
2007	356	94	219	0	55	4
2008	313	524	257	30	47	33
2009	210	835	246	333	58	130
2010	207	951	213	787	40	210
2011	9	(1040)	95	(945)	0	(256)
Total all years	**6756**	**3645**	**5764**	**2096**	**41**	**623**

Selection intensity increases by birth year

Selected Nordic bulls VikingGenetics Proportion selected

Birth year	HOL	RDC	Jersey
2007	80%	100%	90%
2008	40%	90%	60%
2009	25%	40%	30%
2010	15%	20%	15%
2011	(15%)	-	-

Significant increasing selection intensity

Superiority in NTM selected bullse

Bulls born 2009

Average NTM	HOL	RDC	Jersey
All	10.3	9.2	7.4
Selected	16.2	12.0	10.6

Superiority of genomic selected young bulls will increase with increasing selection intensity

Superiority selected Holstein bulls

Bulls born 2008-2011

	Protein yield		Difference
Birth year	Selected	All	
2008	111.6	107.9	3.7
2009	115.1	110.1	5.0
2010	115.8	112.3	3.5
2011	118.1	112.7	5.4

Superiority selected Holstein bulls

Bulls born 2008-2011

	Udder health		Difference
Birth year	Selected	All	
2008	102.7	100.0	2.7
2009	103.7	99.8	3.9
2010	107.3	102.3	5.0
2011	106.1	102.3	3.8

Results – Top bulls – Nov11

	HOL[1]	RDC[1]	JER[2]
% Young bulls >20mth	71	54	50
% Progeny tested	29	46	50
Min NTM	+26	+20	+20

[1] Top50

[2] Top20

Results – Top bulls Nov11 (Aug11)

	HOL[1]	RDC[1]	JER[2]
% Young bulls >20mth	71 (60)	54 (37)	50 (27)
% Progeny tested	29 (40)	46 (63)	50 (73)
Min NTM	+26 (+24)	+20 (+20)	+20 (+17)

[1] Top50

[2] Top20

Frequency of young bulls at the top list dependent on:

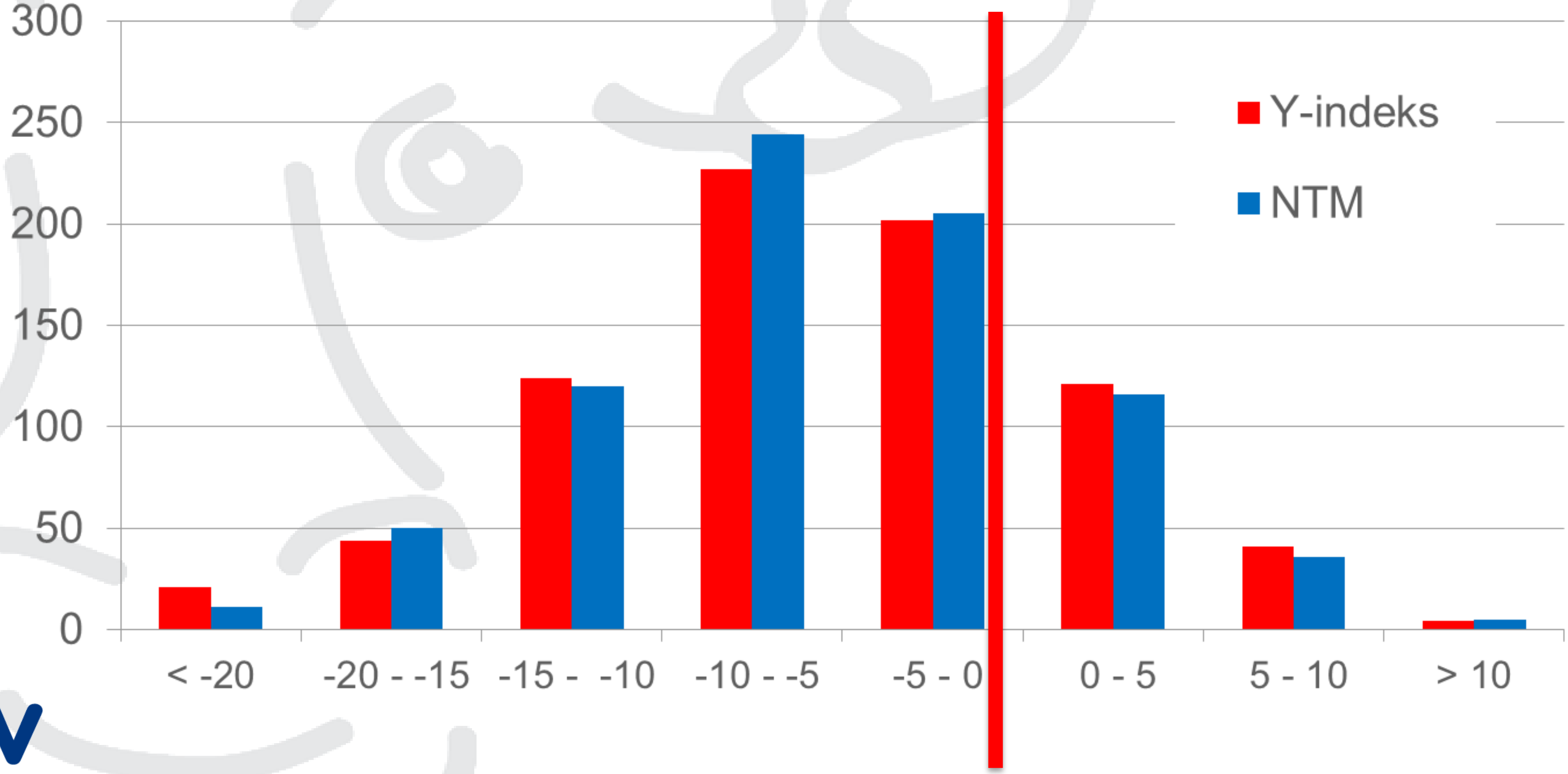
- Number og genotyped young bulls compared to number of traditional progeny tested bulls. If a lot of young bulls are genotyped the proportion of young bulls will increase (change Aug to Nov).
- The birth years included – are all bull calves included or only bull calves > 20 month. Including the very young bull calves increase the proportion of young bulls

Proportion of young bulls on the top list:

- CDN, FRA, NLD, DEU and the Nordic countries have all about 60-70% young bulls at the top list

- Proportion of young bulls on the top list are somewhat higher in US. US includes the bull dam information in the GEBV for young bulls. Americans talk about that that young bulls are over evaluated (EAAP 2011)

Changes in EBV (with and with out genomic information) for genotyped Holstein heifers



Next steps in routine estimation of GEBVs

- NRF in RDC reference
  - Goal 15 Dec 2011 – test and comparison in November 2011 – increase in reliability 1-2%
- Reliability – goal to give reliabilities also to GEBVs
  - Bulls without a sire/MGS in ref. pop. Su&Lund (2009)
- Claw health - test possibilities based on current data
- Genotyped cows and reference bulls
  - Goal to combine genomic information + own performance (ongoing)
- Imputation

Models for future routine genomic prediction

	Phenotype	Genotype	Steps	Method
Today	DRP-sire	50K	2 step	GBLUP
Future A	DRP-sire DRP-AM	50K + 800K + 6k	2 step	GBLUP Bayes
Future B	DRP-sire DRP-AM Phenotypes-AM	50K + 800K + 6k	1 step	GBLUP Bayes

More info in next presentation