ICAR meeting 2009

State of the art of genomic evaluation
and international implications

Reinhard Reents

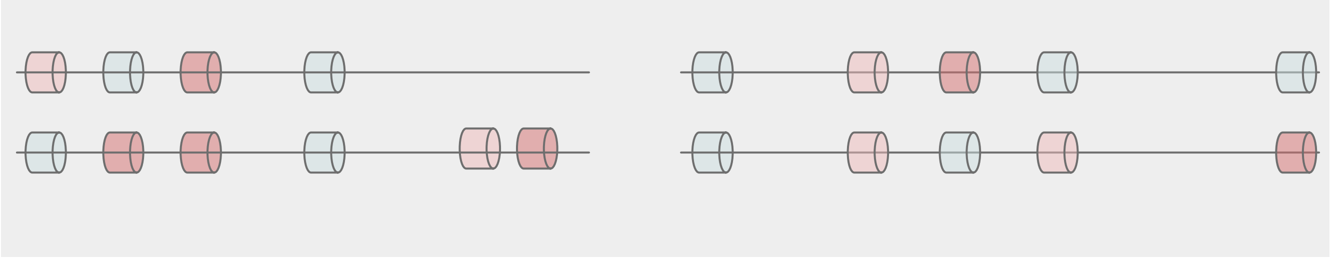
vit, IT solutions for Animal Prduction
Chairman of the Steering Committee of Interbull

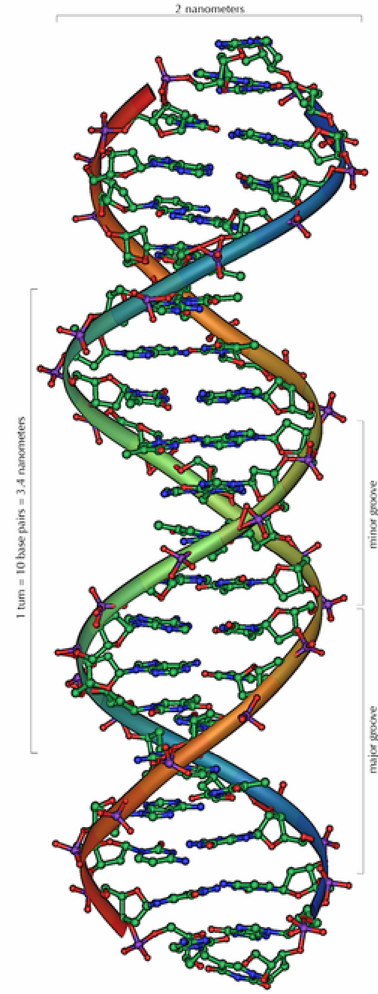
Knowledge about genomics

- Watson und Crick 1953 Model of double helix →

  (structure of DNA where two chains are held together by hydrogen bonds in a double spiral)

- But no information about effect of individual genes/animals:





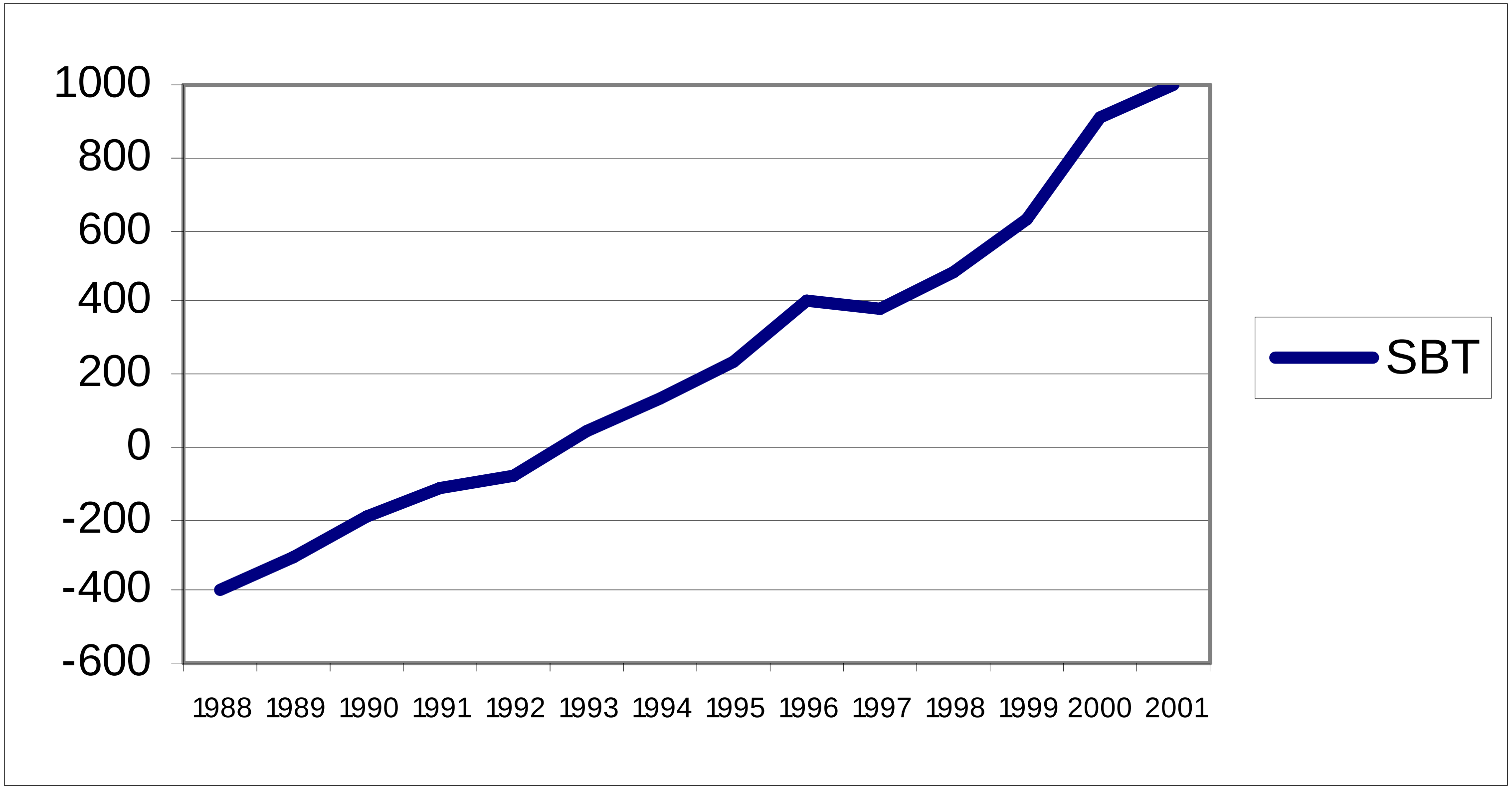
Selection from about 1960 to ~ 2000

- Quantitativ-genetic concepts
    - (Wright, Lush, Henderson)
      -> additiv genetic model
- Genetic evaluation
    - Separate phenotypic observations (eg 9850 kg milk) in
        - additiv genetic effect → estimated breeding value (eg. + 1430 kg M)
        - Systematic environmental effect
        - Residual effect
- Ranking based on estimated breeding values (EBVs)

Selection based on EBVs

- Necessary elements
  - Phenotypic observations
    - Milk yield, somatic cell counts, type traits, etc.
  - Pedigree data
  - Data structure (across herds/environments)
    - Artificial insemination gives optimal structure to estimate EBVs that rank the animals best and unbiased in many environments
  - Algorithms (Henderson, Schaeffer&Kennedy, Misztal, etc) and computing power
    - ➔ BLUP methodology, which result in highly reliable EBVs (85-99%) for bulls with a **progeny test of 100-150 daughters**
    - ➔ Transformation of these EBVs since 15 years via Interbull MACE
      - ➔ **Bulls that are marketed worldwide**

Genetic trend in German Holstein bulls



but

- Genetic gain / costs
  - High generation interval due to progeny test
    - ➔Expensive
    - ➔Genetic gain per year not very high
  - Reliability of a pedigree index (=0,5 EBV sire + 0.5 EBV dam) is low (25 – 35%)
  - Reliability of a cow EBV < lower as r^2 of a bull EBV

- Aim:
  - Increase of reliability of young animals

- Solution → use of genomic informationn

SNP - genotyping

- SNP = Single-nucleotide polymorphism
- Genotype = Which alleles of the nucleotides **A-T,C-G** an animal carries
- Genome = contains 3 billion base pairs
- Ca. 50.000 SNPs at a cost of about 200 EUR

		Genotype**:**
Tier 1:	...AGGCACC GCAATCCACG GAGGC**T**ACGC CCTCACCGGA GGTTTCGCTC TCCACGG... ...AGGCACC GCAATCCACG GAGGC**T**ACGC CCTCACCGGA GGTTTCGCTC TCCACGG...	**TT**
Tier 2:	...AGGCACC GCAATCCACG GAGGC**A**ACGC CCTCACCGGA GGTTTCGCTC TCCACGG... ...AGGCACC GCAATCCACG GAGGC**A**ACGC CCTCACCGGA GGTTTCGCTC TCCACGG...	**AA**
Tier 3:	...AGGCACC GCAATCCACG GAGGC**T**ACGC CCTCACCGGA GGTTTCGCTC TCCACGG... ...AGGCACC GCAATCCACG GAGGC**A**ACGC CCTCACCGGA GGTTTCGCTC TCCACGG...	**AT**
Tier n:	...AGGCACC GCAATCCACG GAGGC**A**ACGC CCTCACCGGA GGTTTCGCTC TCCACGG... ...AGGCACC GCAATCCACG GAGGC**A**ACGC CCTCACCGGA GGTTTCGCTC TCCACGG...	**AA**

Eg position on Chromosom 6 # 43.675.239

Illumina cattle SNP-Chip: BovineSNP50 Beadchip

Number of SNPs:	54.001
Eg informative in Holsteins:	42.730
Ø distance between 2 SNPs (n base-pairs):	51.500
Minimum distance (n base pairs):	20.000
Amount of DNA needed:	> 2 µg
eg. 1 ml blood contains	~ 200 µg
1 dose of sperm contains	~ 50 µg
(10 roots of a hair contain	~ 1 µg)

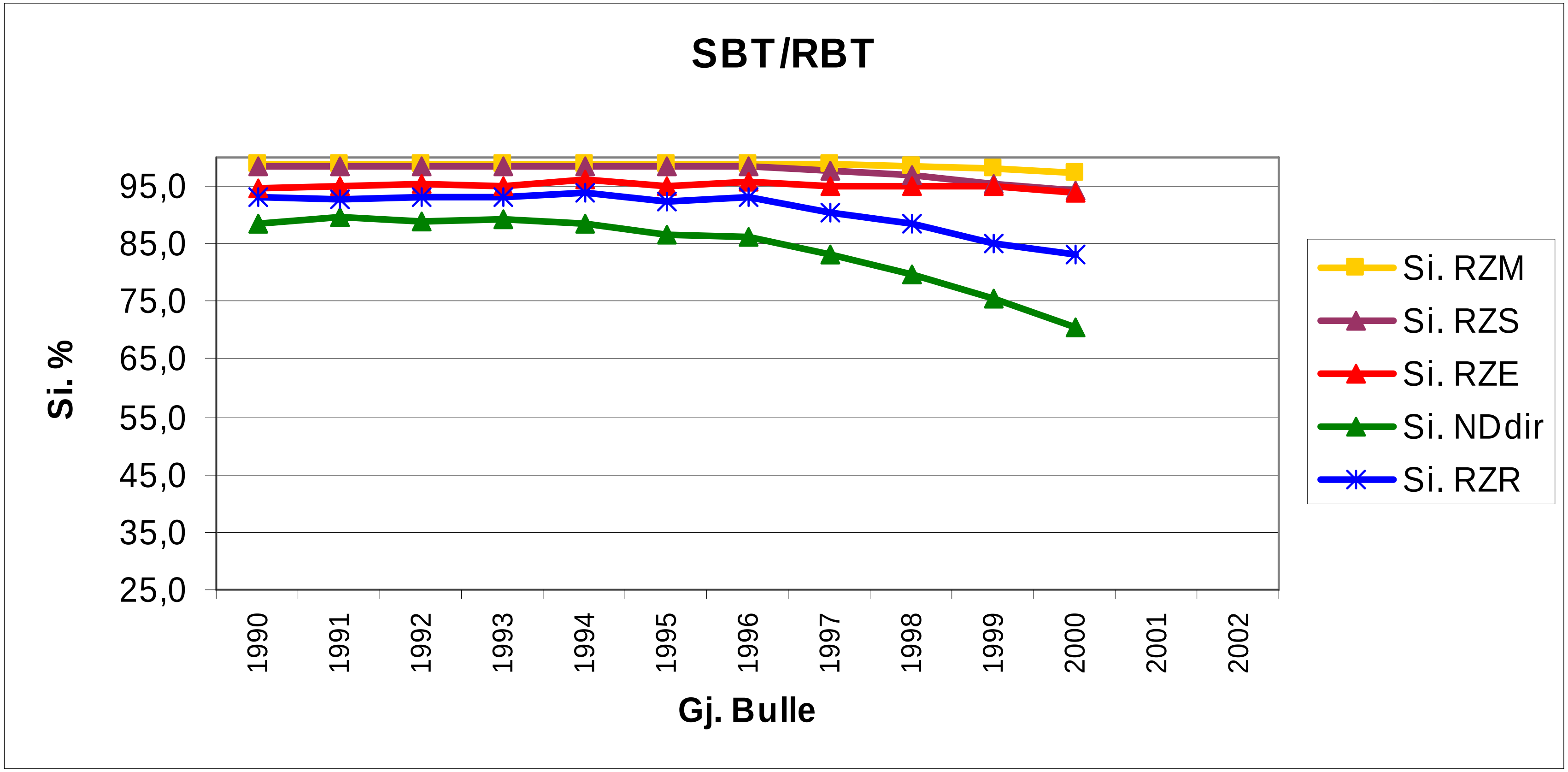
Use in practical application

- Lab is an important part, BUT
  SNP data has per se no information on ‚traits‘

Steps:

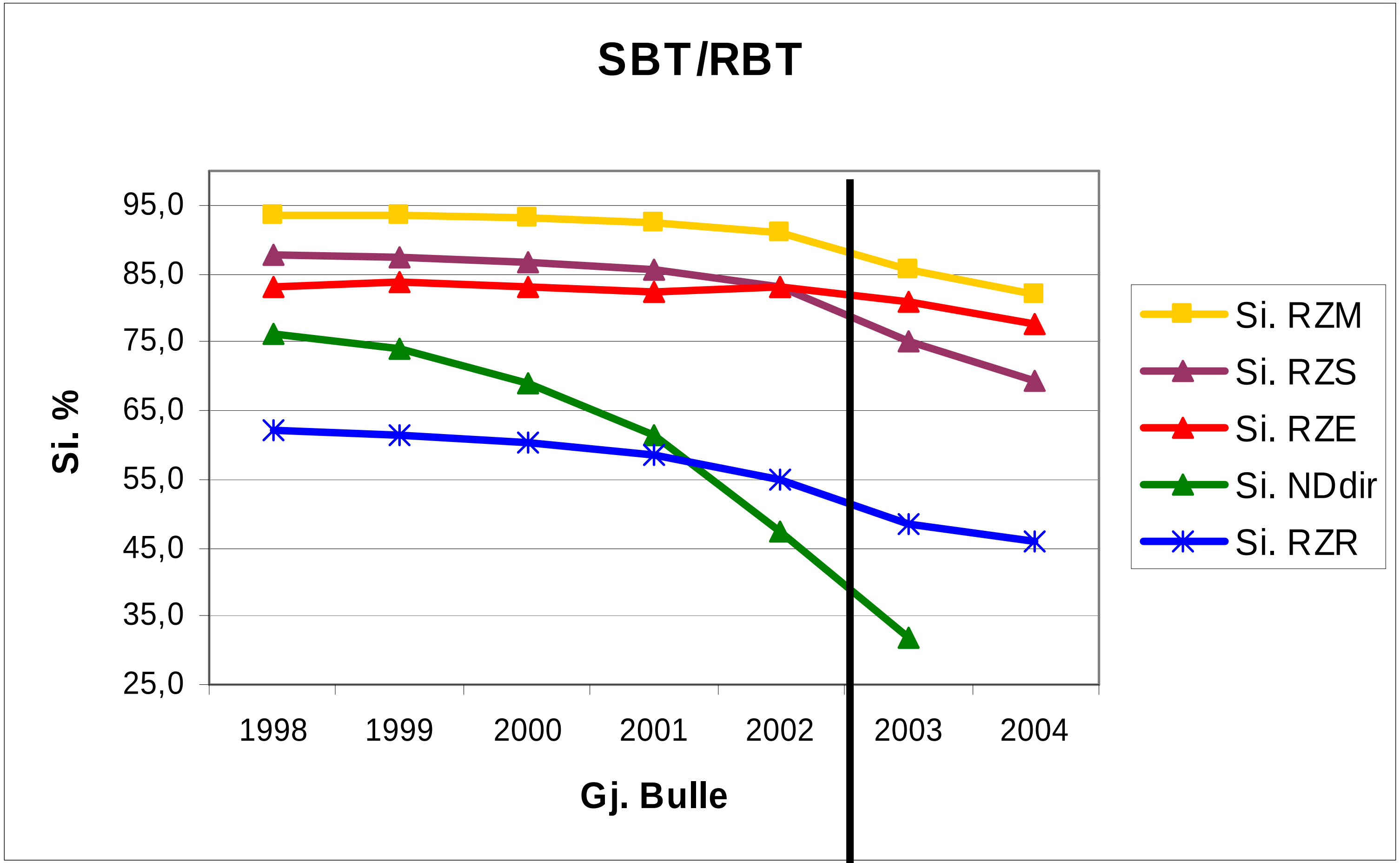
- Genotype animals that have reliable EBVs from ‚conventional‘ genetic evaluation
- Calculate regression formulas so that SNPs explain well the conventional EBV
- Use the regression formulas derived by historic data to evaluate young animals
- Select among these young animals

Available data, r² for different traits in German Holsteins (n= ~ 700, > 500 daughters)



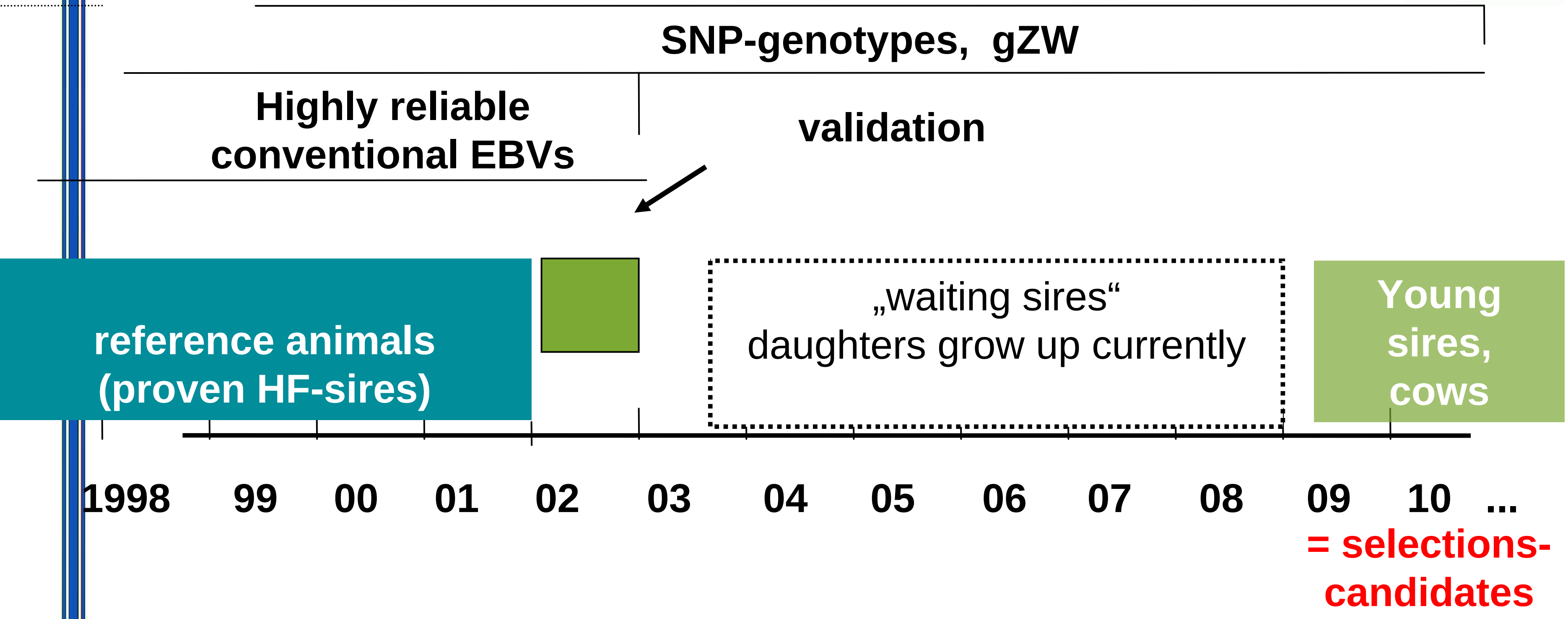
n WE-Bullen	126	130	100	101	91	77	77	101	72	46	16	1	

R^2 for progeny tested bulls German Holstein bulls (850/year) ca. 115 daughters per bull



n M/S/E	829	853	872	1010	1014	847	62
n NDdir	829	853	872	1010	991	311	1

Example for application – German Genotrack project; status quo 2009



SNP-genotypes, gZW

Highly reliable conventional EBVs

validation

reference animals (proven HF-sires)

„waiting sires“ daughters grow up currently

Young sires, cows

1998 99 00 01 02 03 04 05 06 07 08 09 10 ...

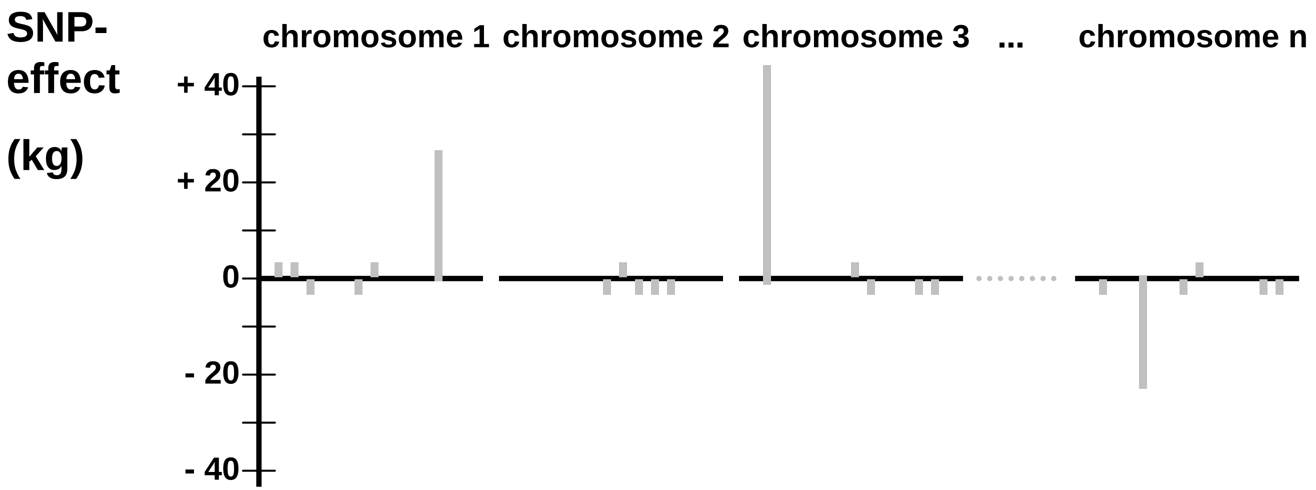
= selections-candidates

- Estimation of SNP-effects: complete age-groups HF → 3500 sires
- Validation of SNP-effects:
  - Sires getting first progeny-based EBV in 2009 → ~ 900 sires
  - Strategy: correlation² between DGV: conventional EBV

Modell genomic evaluation – quality of results

- Quality of phenotypic data: conventional breeding value estimation
  - ➔ Importance of conventional EBV

- Estimation of SNP-effetcs using „*informative*“ reference group
  - The more sires, the more accurate SNP-effects estimates
- Application of SNP-effects on young animals



1 + 1 - 1 - 1 + 1 + 25 - 1 + 1 - 1 - 1 - 1 + 42 + 1 - 1 - 1 - 1 - 1 - .... 22 - 1 + 1 - 1 - 1 = **+38**

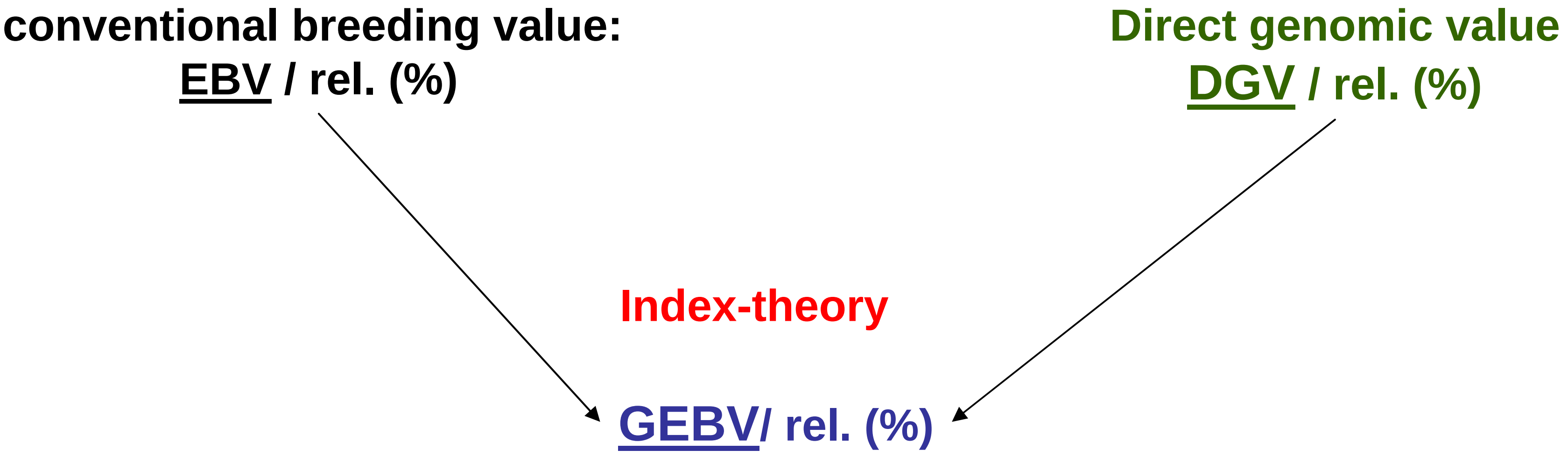
→ **DGV**

Size of reference sample, van Raden, IB meeting, June 2008

Gain in reliability over PA in US

Bulls		R^2 for Net Merit		
Predictor	Predictee	PA	Genomic	Gain
1151	251	8	12	4
2130	261	8	17	9
2609	510	8	21	13
3576	1759	11	28	17

Combination: conventional EBVs + genomic EBVs



conventional breeding value:
EBV / rel. (%)

Direct genomic value
DGV / rel. (%)

Index-theory

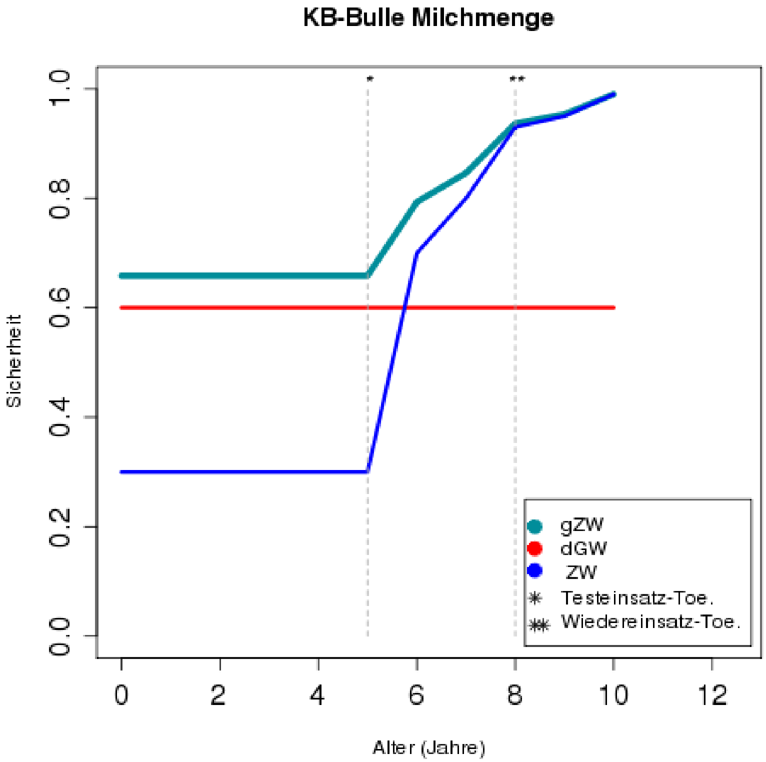
GEBV/ rel. (%)

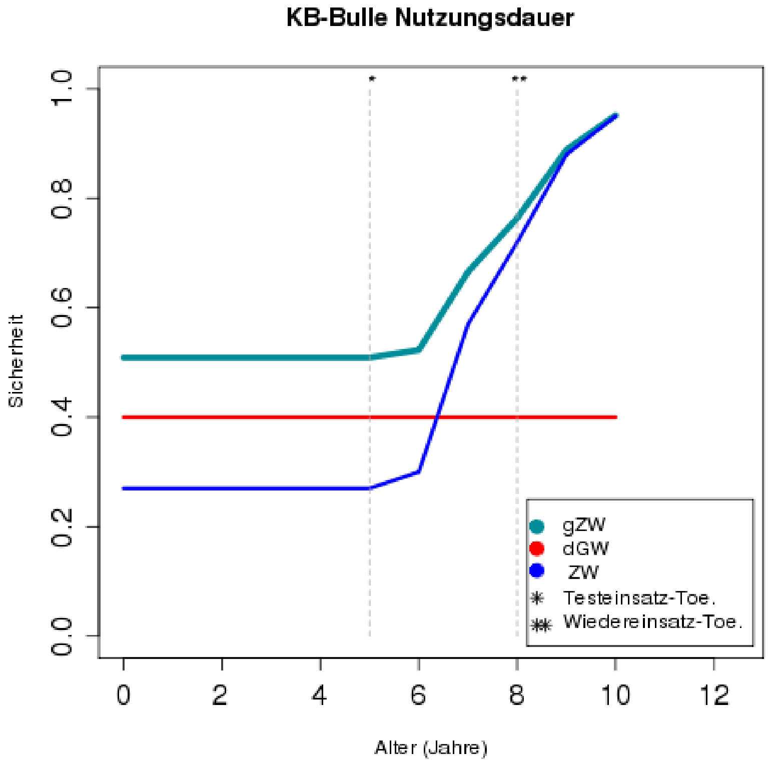
combined **genomic breeding value**

Open questions:

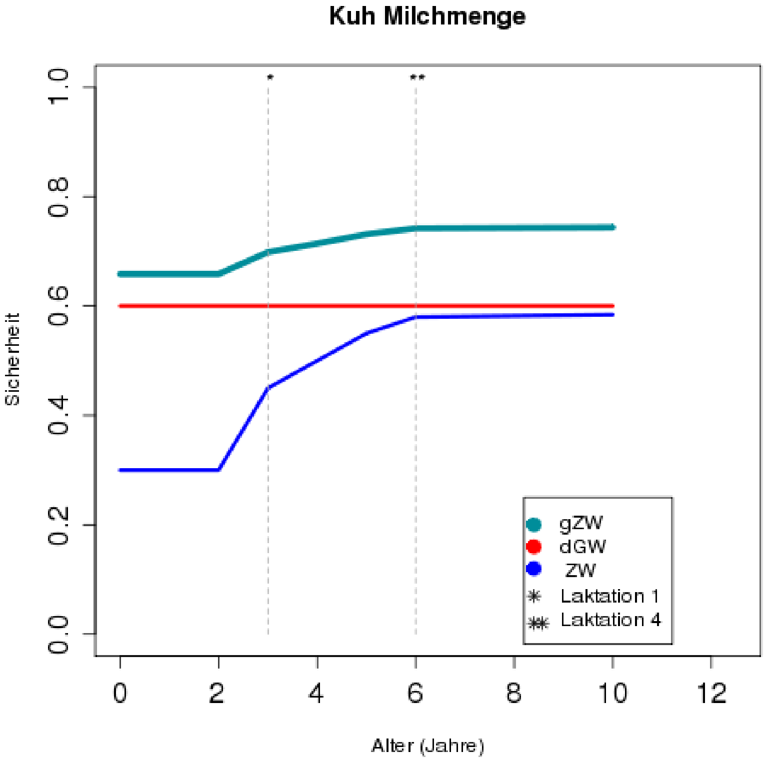
- Realistic reliability of DGV ?
- If and how to combine DGV with conventional pedigree EBVs?

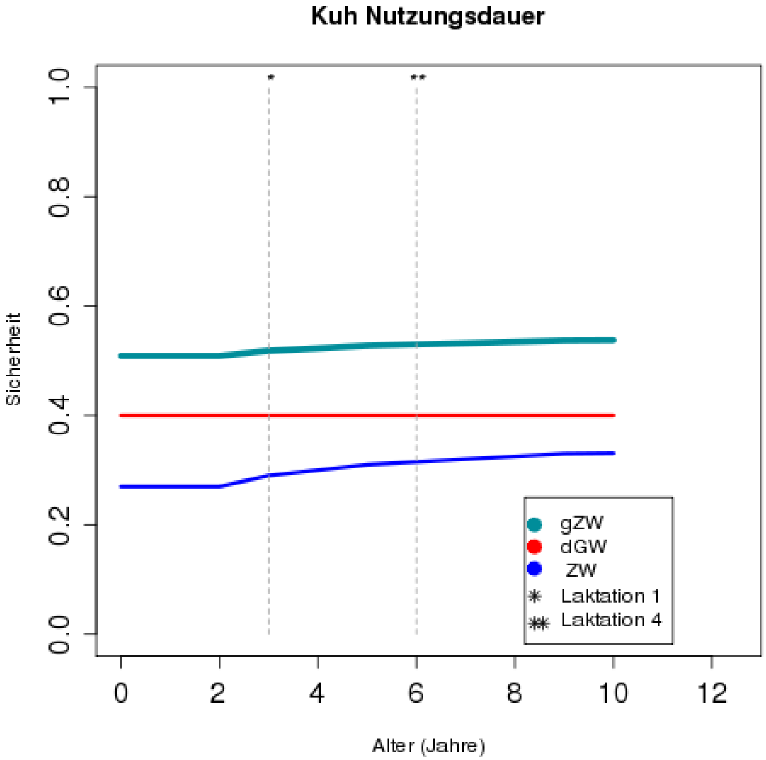
Increase in reliability of the BVs: AI bulls


KB-Bulle Milchmenge


KB-Bulle Nutzungsdauer

Increase in reliability of the BVs: cows


Kuh Milchmenge


Kuh Nutzungsdauer

Spectrum for practical application

1. Selection of young sires for A.I.
   - = selection of young sires based on GEBV
   - Conventional progeny test follows

2. Selection of bull dams
   - = selection & mating of bull dams based on DGV/GEBV

3. Selection of sires in A.I.
   - = selection of sires in A.I. based on GEBV
   - Sires age: 1 year, semen unlimited available
   - **Bulls without a progeny test → acceptance by farmers?**
   - **Increased use of natural service sires that do have a GEBV instead of AI bulls?**

4. Selection of sire of sons
   - Sire of sons genomic proven without progeny test

5. Selection of cows
   - Realistic if low density SNP-chip (cost-effective) available
     genotype – environment interaction???

Status quo on country level: genomic selection (Jan 2009)

country	Project started	chip	Size of reference population	GEBV	Official Implementation	Internal Implementation
USA	2003	Illumina	4.422 sires + 947 cows	✓	Jan. 2009	2008
CDN	2003	Illumina	4.127	✓	April 2009 ?	?
NZL	2005	Illumina	1450	✓	2010 ?	(Aug. 2008)
AUS	2005	Illumina	1.600	✓(?)	2010 ?	
NLD	2005	MG1 / MG2	1.500	✓	2009 ?	Aug. 2008
FRA	2007	Illumina	1.750	(MAS+ gZW)	?	Fall 2008
DFS	2007	Illumina	2.012	?	?	Aug. 2008
(DEU /HOL)	2008	Illumina	3.000	✓	2010	2009

Summary

- **Genomic evaluation is a very useful additional tool, but it can only be used if**
    - **Large amount of phenotypic data is collected on the traits of interest**
    - **Reliable pedigree information**
    - **Sophisticated conventional genomic evaluation**

- **Potential of genomic evaluation can only be gathered if MORE phenotypic data is collected**
    - **Functional traits (well defined eg by the ICAR WG ‚Functional Traits‘)**
    - **New traits → e.g. composition of milk**

- **Many options how to use this tool in breeding programes have still to be examined**

Discussion on the methodology

Interbull / ICAR meeting 27.1. - 30.1.2009

- Interbull workshop (27. - 28.1.2009, 101 participants)
  - Report Task Force (→ 7 experts from genetic evaluation centres)
  - Summary of the questionnaire
  - Conclusions and recommendations from the discussion groups

- Interbull Steering Committee (29.1.)
- ICAR Ex. Board (30.1.)

Results Interbull Workshop 27.1. - 28.1.2009

- Report Task Force Genomic Evaluation
  - Methodology works
  - but: assignment of an unbiased reliability measure for the genomic EBVs is not achieved yet → clear tendency to overestimate the r^2
    - Approximation of r^2:
      use prediction formula for group of bulls with genotypes and conventional EBVs, but were **not** part of the reference population
      → basis of validation procedure

- Results from the questionnaire
  - 20 countries use → Illumina Bovine SNP50 BeadChip
  - Holland → Customized Illumina 60K BeadChip
  - Norway → Affymetrix

Participants identified need for new Interbull services

All countries

- Express the wish to share knowledge and improve the methodology within the Interbull framework
  - Documentation of
    - Description of the methodology applied
    - Data used, publication rules within country, etc.

    → To be published by Interbull like methods for conventional genetic evaluation

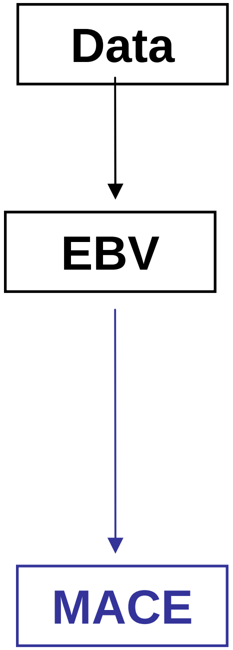
- Urgent need for Interbull validation procedure
  - Genomic evaluation system → unbiased DGV and GEBV
  - Unbiased r^2 for DGV and GEBV

→ Both relevant for approval within EC

Specific Interbull services:
a) countries with own genome programe

- Urgent need for services
  - Conversion of DGV and/or GEBV on other countries scales
  - GMACE
  - Logistic for exchange of SNP genotypic data within the Interbull framework

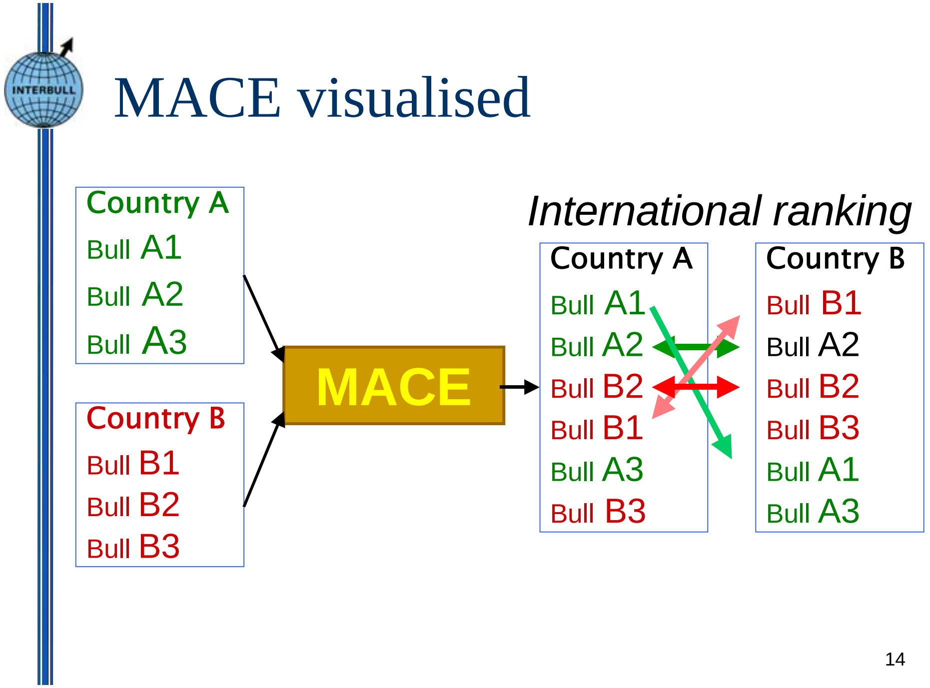
MACE



Data

EBV

MACE



MACE visualised

Country A
Bull A1
Bull A2
Bull A3

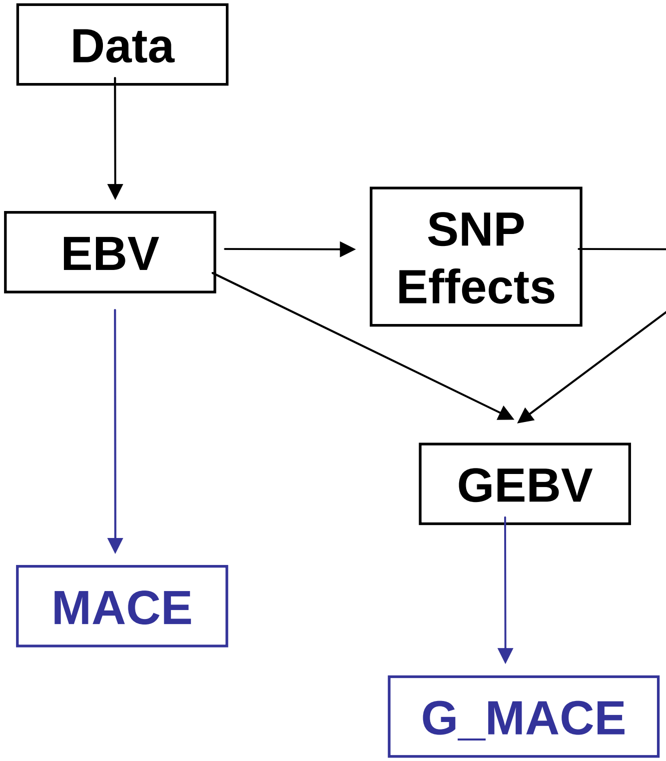
Country B
Bull B1
Bull B2
Bull B3

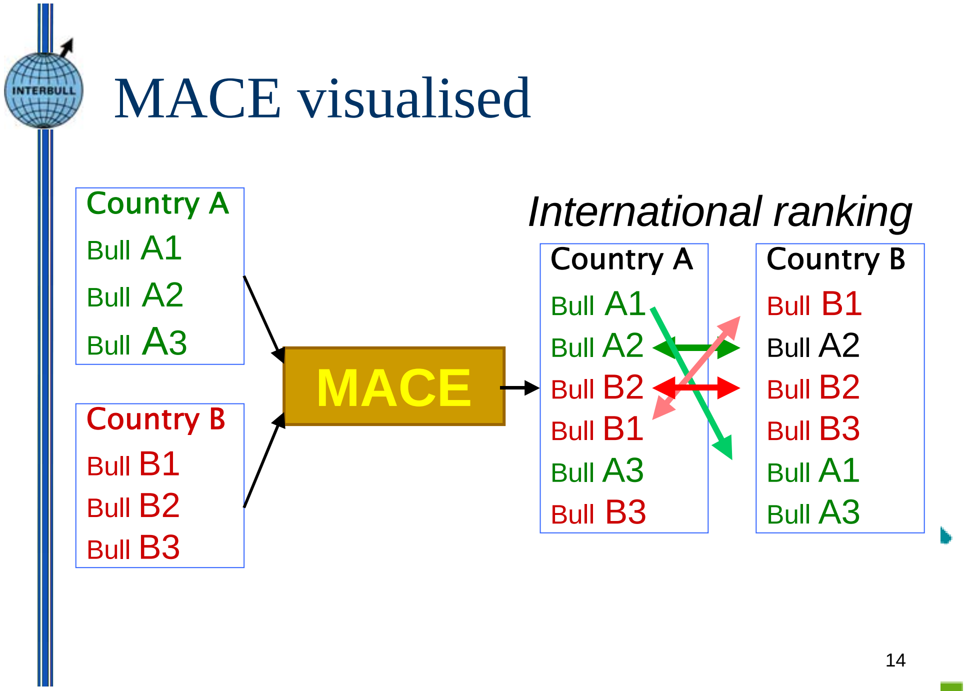
MACE

International ranking

Country A	Country B
Bull A1	Bull B1
Bull A2	Bull A2
Bull B2	Bull B2
Bull B1	Bull B3
Bull A3	Bull A1
Bull B3	Bull A3

MACE → GMACE





MACE visualised

Country A
Bull A1
Bull A2
Bull A3

Country B
Bull B1
Bull B2
Bull B3

MACE

International ranking

Country A	Country B
Bull A1	Bull B1
Bull A2	Bull A2
Bull B2	Bull B2
Bull B1	Bull B3
Bull A3	Bull A1
Bull B3	Bull A3

Specific Interbull services:
b) countries without an own genome programe

- Seek implementation directly within Interbull sharing data from several countries/populations
  - Increase size of reference sample
  - Decrease costs for technical implementation
    - IB hosts genomic SNP data
    - IB derives prediction formulas
    - Calculates DGV for these countries

      Interest from
      - ‘smaller’ european countries → Holstein
      - Brown Swiss breeders from several countries

Interbull Steering committee decisions

- Recognition of neutral, unbiased statistics from Interbull has to be maintained
  - Any recommendation for use of DGV/GEBV can only be granted after the validation test has been passed

- Validation test shall be developed by the IB Task Force until summer 2009
  - Approval at the Interbull Barcelona meeting in August 2009

- Steps for implementation
  - A) use of Interbull conversion formulas (summer/fall 2009)
    B) GMACE (beginning of 2010)
    C) develop a plattform that allows exchange of SNP 0/1 data for pairs of countries (2010/2011)
    D) a position of a PostDoc for implementation of this technology is to be filled

ICAR board

- Results of the Interbull workshop will be summarised and transmitted to the EC
- No use of DGV/GEBV with the label of Interbull can be done before a successful validation
  (→ 2nd half of 2009)
- On an interim basis DGV/GEBV have to be labeled

Summary

- Significant impovement of the methodology during the last 12 months
- Very successful meeting in Uppsala in January 2009
- Large breeding programes want to cooperate in several areas
  - Improvement of the methodology
  - Sharing genotypes of the reference sample
- Small countries / populations seek implementation of the complete system within Interbull centre
- Interbull has made a strategic plan to introduce this new technology into its portfolio

Thank you for your attention