Applications of Genomics –

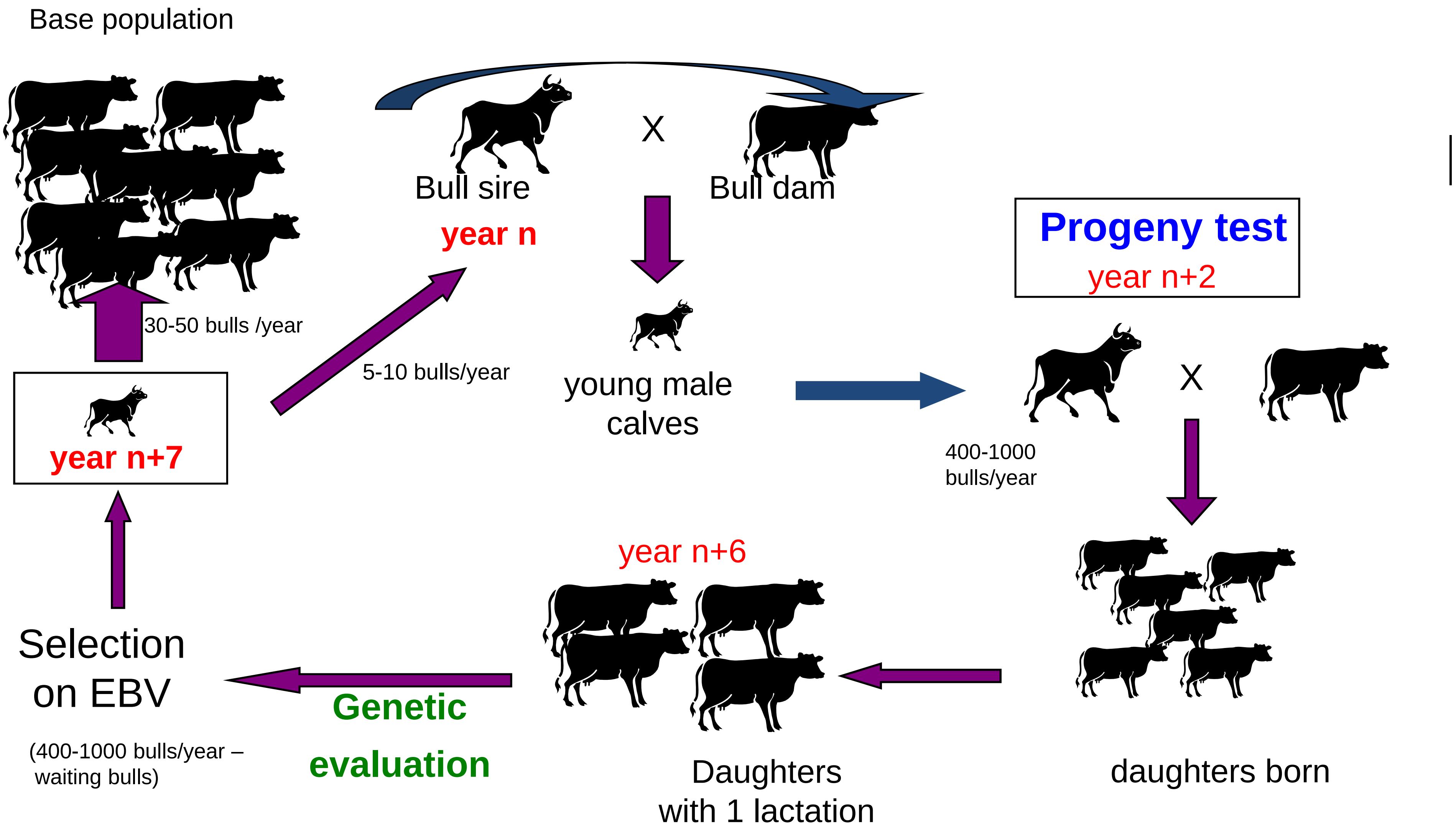
Introduction of genomics selection procedures in existing genetic improvement programmes

25-26 August, **2014**

Animal Breeding Group
NDDB
Anand

Genomic Selection - Concept

A classical progeny testing scheme



Base population

30-50 bulls /year

year n+7

Bull sire

year n

5-10 bulls/year

X

Bull dam

young male
calves

Progeny test

year n+2

X

400-1000
bulls/year

year n+6

Selection
on EBV

(400-1000 bulls/year –
waiting bulls)

Genetic

evaluation

Daughters
with 1 lactation

daughters born

Model used for Breeding Value Estimation – Animal BLUP

$$y = Xb + Za + e$$

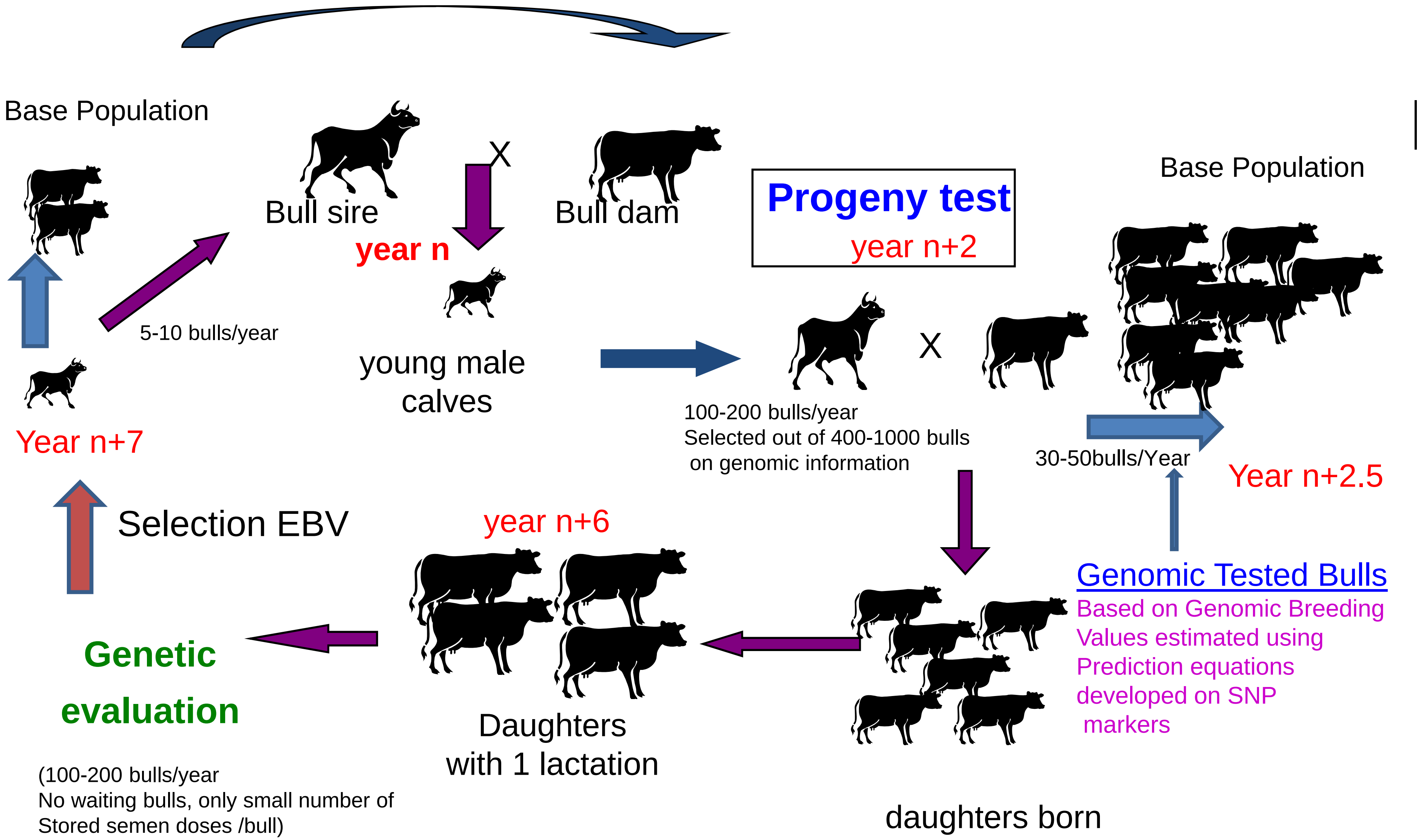
Where:

- y = n x 1 vector of observations; n= number of records
- b= p x 1 vector of fixed effects; p=number of levels for fixed effects
- a= q x 1 vector of random animal effects; q= number of levels of random effects
- e= n x1 vector of random residual effects
- X= design matrix of order n x p, which relates records to fixed effects
- Z= design matrix of order n x q, which relates records to random animal effects

$$\begin{bmatrix} \hat{b} \\ \hat{a} \end{bmatrix} = \begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + A^{-1}\alpha \end{bmatrix}^{-1} \begin{bmatrix} X'y \\ Z'y \end{bmatrix}$$

A is the relational matrix based on pedigree data and $\alpha = \sigma^2_e / \sigma^2_a$

Genomic Tested Young Sire Model – World over



Base Population

Bull sire

X

year n

Bull dam

5-10 bulls/year

young male calves

Progeny test

year n+2

Base Population

100-200 bulls/year
Selected out of 400-1000 bulls
on genomic information

X

30-50bulls/Year

Year n+2.5

Year n+7

Selection EBV

year n+6

Genomic Tested Bulls

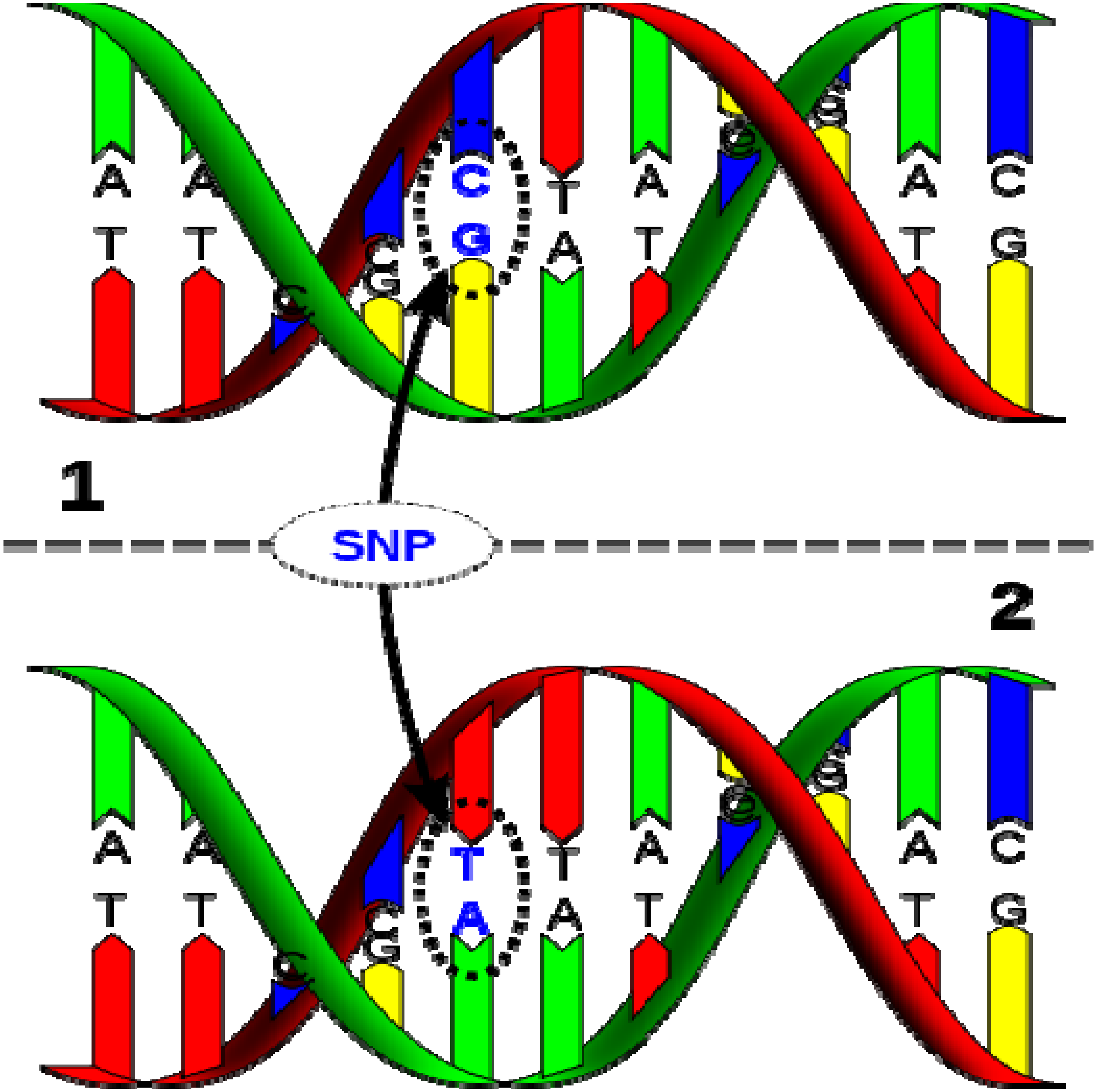
Based on Genomic Breeding
Values estimated using
Prediction equations
developed on SNP
markers

Genetic evaluation

(100-200 bulls/year
No waiting bulls, only small number of
Stored semen doses /bull)

Daughters
with 1 lactation

daughters born


A A C C T A G A C
T T C G G A T G T G
1
SNP
2
A A C T T A G A C
T T C G A A T G T G

Genomic Selection-Basics

- Bulls having breeding value based on daughter performance and genotyped data is referred to as reference population
- Such Bulls are divided in to two parts – Training set and Validation set
- Genotyping of bulls DNA is done at SNP level
- Using SNP data and breeding values of training set, the effect of SNPs are estimated on traits and genomic breeding value prediction equations are developed.

Genomic Selection

- These equations are validated in validation population
- Predicted breeding values based on equations are compared with actual breeding value based on progeny test results for the validation set
- Models are adjusted to get maximum correlations
- Future breeding values of Young bulls' are predicted by these equations based on genotype information (SNPs) alone.

Methods for Genomic Breeding value estimation

Two approches have been tried:

(i) Direct estimation of marker effects using Bayes B approach

(ii) gBLUP using genomic relation matrix derived from markers among all individuals

Direct estimation of marker effects

$$Y_i = u + \sum^k_{j=1} X_{ij} B_j \partial_j + e_i$$

Where:

Yi is the phenotype of animal I

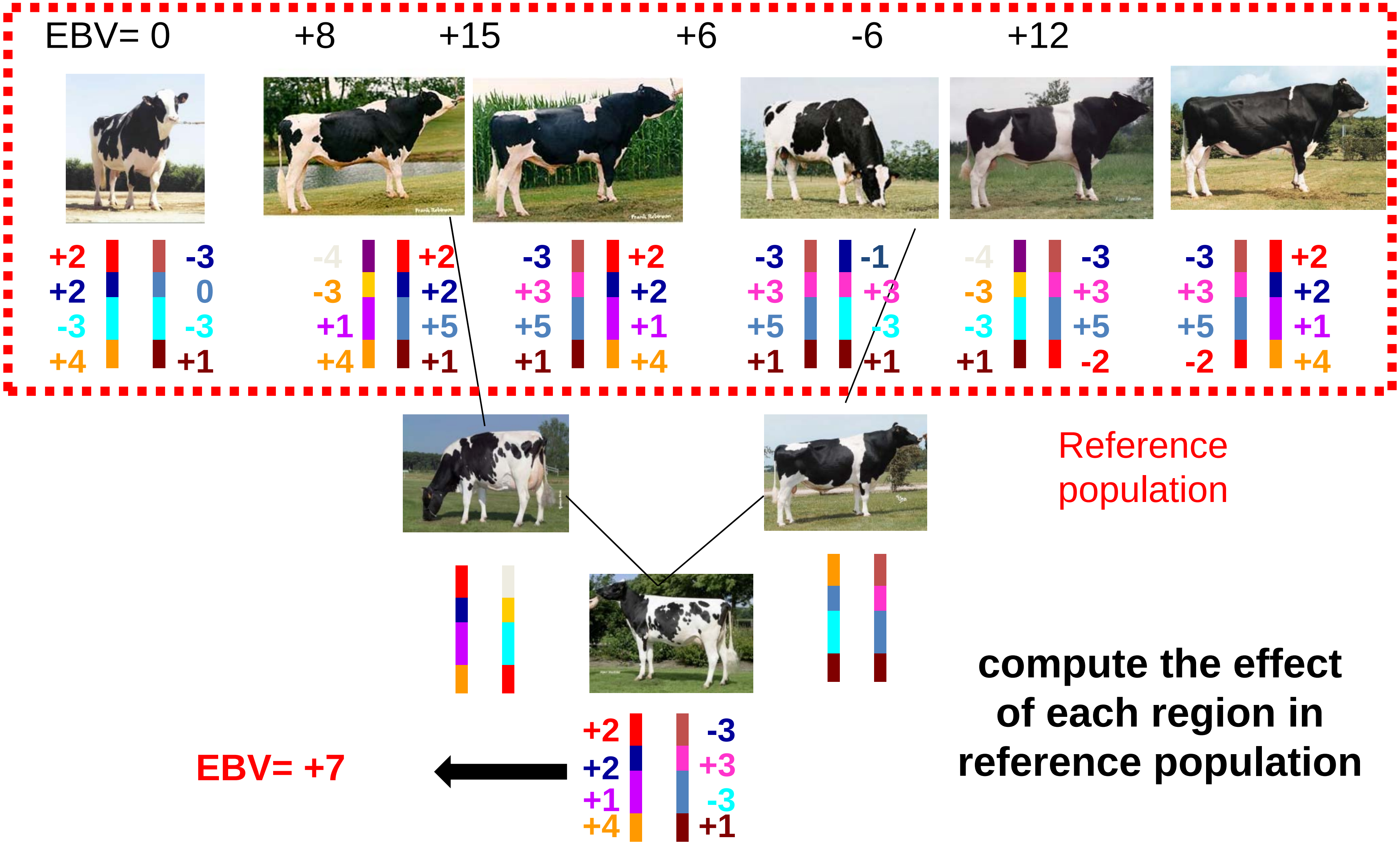
U is the overall mean

K is the number of marker loci

Xij is the genotype at locus j which is coded as 0, 1, or 2

Bj is the allel substitution effect at locus j,
∂j is a 0/1 variable indicating the absence or presence of locus j in the model, ei is the random residual effect

What is genomic selection?



Reference population

compute the effect of each region in reference population

EBV= +7

Breeding Value Estimation – through gBLUP

$$y = Xb + Za + e$$

Where:

y = n x 1 vector of observations; n= number of records

b= p x 1 vector of fixed effects; p=number of levels for fixed effects

a= q x 1 vector of random animal effects; q= number of levels of random effects

e= n x1 vector of random residual effects

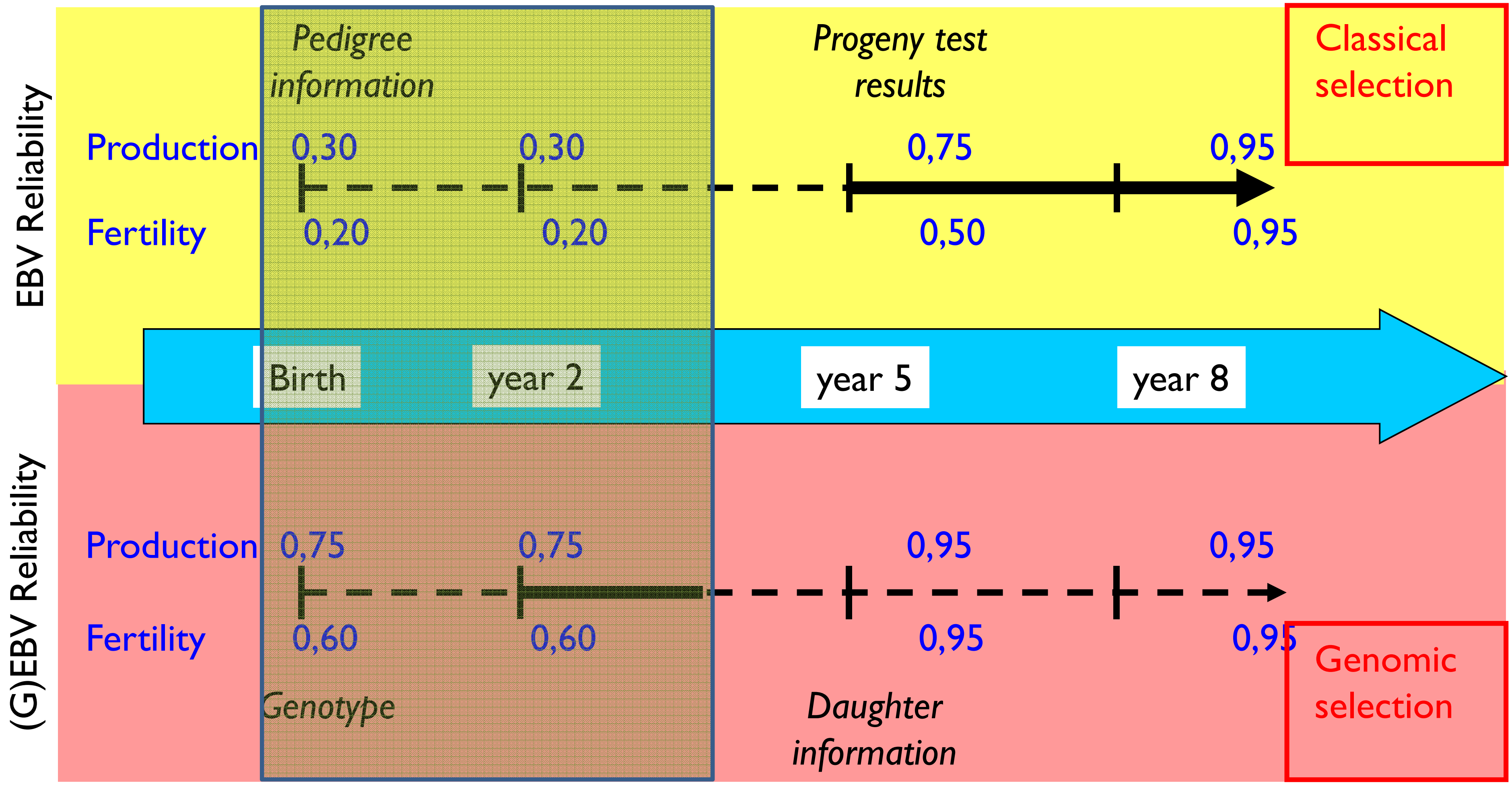
X= design matrix of order n x p, which relates records to fixed effects

Z= design matrix of order n x q, which relates records to random animal effects

$$\begin{bmatrix} \hat{b} \\ \hat{a} \end{bmatrix} = \begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + G^{-1}\alpha \end{bmatrix}^{-1} \begin{bmatrix} X'y \\ Z'y \end{bmatrix}$$

G is the relational matrix derived based on marker data and $\alpha = \sigma^2_e / \sigma^2_a$

Consequence



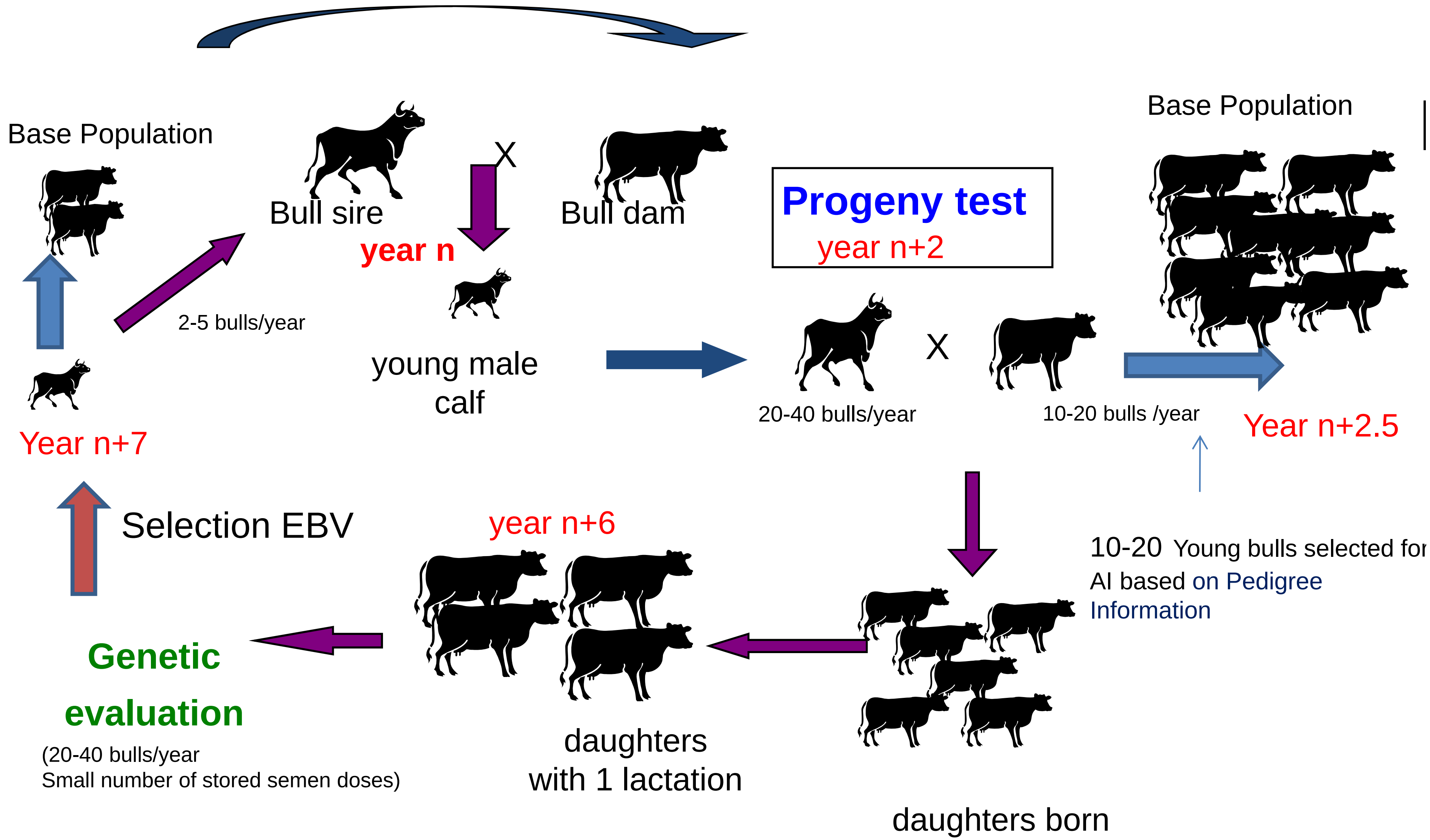
vincent.ducrocq@jouy.inra.fr

Genomic selection –Aftermath across world

- Young bulls selected just based on genomic breeding value and used for AI without waiting for their progeny test results
- Strengthening female nucleus herds and making use of ET and IVF extensively
- Farmers have started using genomic information for selection of heifers
- However they continue to record phenotype - this gives refinement in selection procedure and trends

What can be done in India to introduce Genomic Selection Procedures

Progeny testing scheme – Young Sire Model



Base Population

Bull sire

X

Bull dam

year n

2-5 bulls/year

young male calf

Progeny test

year n+2

Base Population

20-40 bulls/year

X

10-20 bulls /year

Year n+2.5

Year n+7

Selection EBV

year n+6

10-20 Young bulls selected for AI based on Pedigree Information

Genetic evaluation

(20-40 bulls/year
Small number of stored semen doses)

daughters with 1 lactation

daughters born

Traits recorded in PT Projects

- Pedigree information
- Milk yield
- Component yield
  - Fat, Protein, Lactose, SNF
- Reproduction traits:
  - AFC, First Inter calving period, days first bred after calving (First lactation), AI per pregnancy for heifers (DPR) etc.
- Type traits:
  - Confirmation traits (16 traits approved by ICAR)

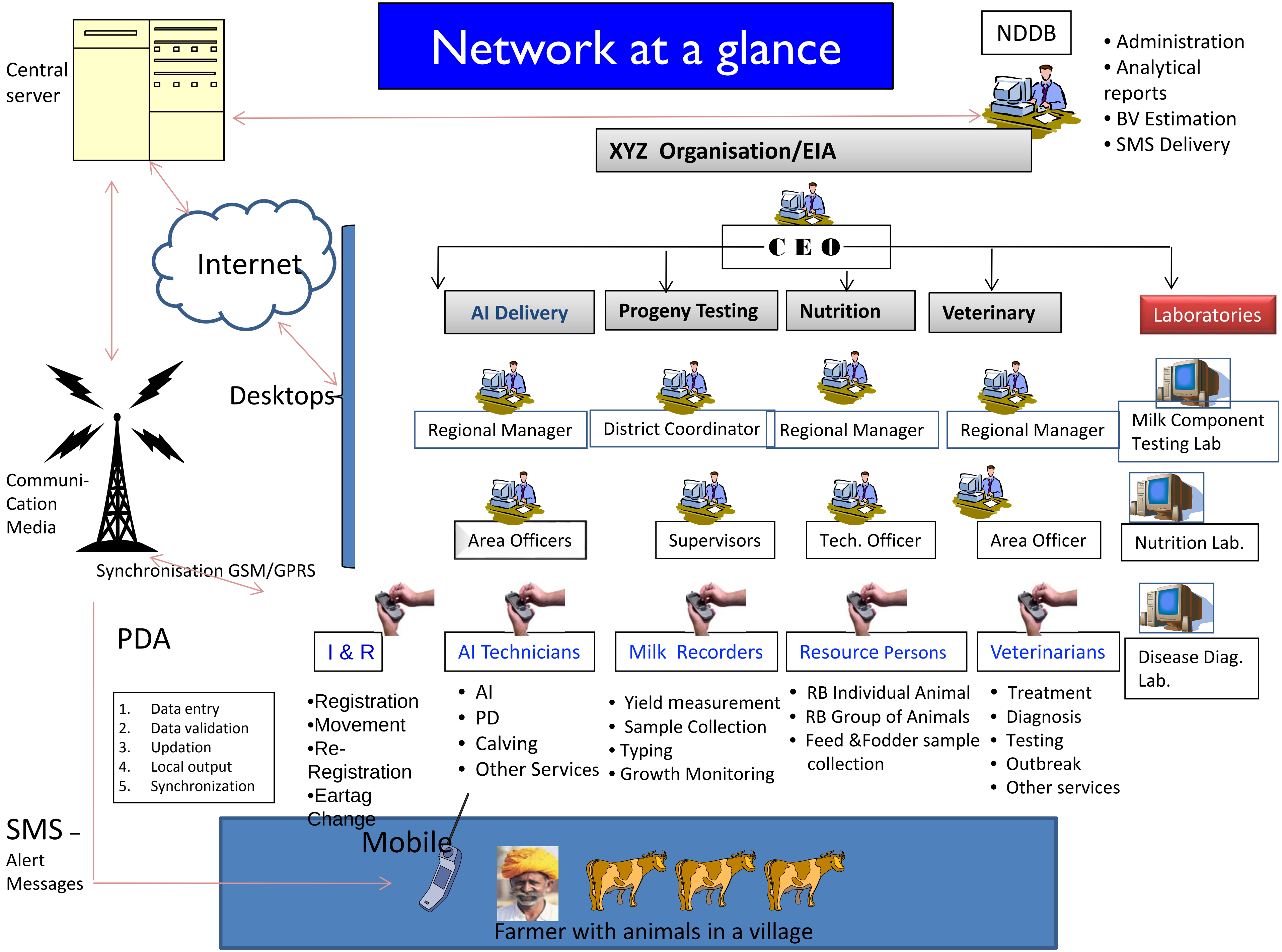
Progeny Testing Projects Approved under NDP-I

Sr. No.	Breeds covered under PT	State (EIA)	Pics
1	HF	Karnataka (KMF).	
2	HF Crossbred	Gujarat (SAGB), Uttarakhand (ULDB) Kerala (KLDB) UP (BAIF).	
3.	Jersey Crossbred	Tamil Nadu (TCMPF), Andhra Pradesh (APLDA).	
4.	Murrah	Haryana, (HLDB) Punjab (PLDB), Gujarat (SAGB), U P (ABRO).	
5.	Mehsana	Gujarat (Mehsana and Banas Unions).	

Current Status of PT Projects Approval

Sr. No.	SPP	Date of approval	Sanctioned amount (Rs. Lakhs)	Timeline
1	HF PT, KMF	27.06.2012	2054.73	2012-18
2	CB Jersey PT, TCMPF	11.09.2012	3928.33	2012-18
3	Murrah PT, SAG	11.09.2012	1061.70	2012-18
4	CB HF PT, SAG	11.09.2012	2175.63	2012-18
5	Mehsana Buffalo PT, Banas	16.11.2012	1211.49	2012-18
6	Mehsana Buffalo PT, Mehsana	16.11.2012	1265.36	2012-18
7	Murrah Buffalo PT, PLDB	18.12.2012	1865.18	2012-18
8	CB Jersey PT, APLDA	11.02.2013	1792.29	2013-18
9	Murrah PT, ABRO	11.03.2013	1501.59	2013-18
10	Murrah PT, HLDB	02.08. 2013	2471.87	2013-18
11	CB HF PT, ULDB	02.08. 2013	1289.84	2013-18
12	CB HF PT, BAIF	25.02.2014	1625.06	2014-18
13	CB HF PT, KLDB	25.02.2014	1485.03	2014-18
Total			**23728.10**	

Network at a glance



Central server

NDDB

- Administration
- Analytical reports
- BV Estimation
- SMS Delivery

XYZ Organisation/EIA

Internet

Desktops

C E O

AI Delivery

Progeny Testing

Nutrition

Veterinary

Laboratories

Regional Manager

District Coordinator

Regional Manager

Regional Manager

Milk Component Testing Lab

Communi-
Cation
Media

Area Officers

Supervisors

Tech. Officer

Area Officer

Nutrition Lab.

Synchronisation GSM/GPRS

PDA

1. Data entry
2. Data validation
3. Updation
4. Local output
5. Synchronisation

I & R

- Registration
- Movement
- Re-Registration
- Eartag Change

AI Technicians

- AI
- PD
- Calving
- Other Services

Milk Recorders

- Yield measurement
- Sample Collection
- Typing
- Growth Monitoring

Resource Persons

- RB Individual Animal
- RB Group of Animals
- Feed &Fodder sample collection

Veterinarians

- Treatment
- Diagnosis
- Testing
- Outbreak
- Other services

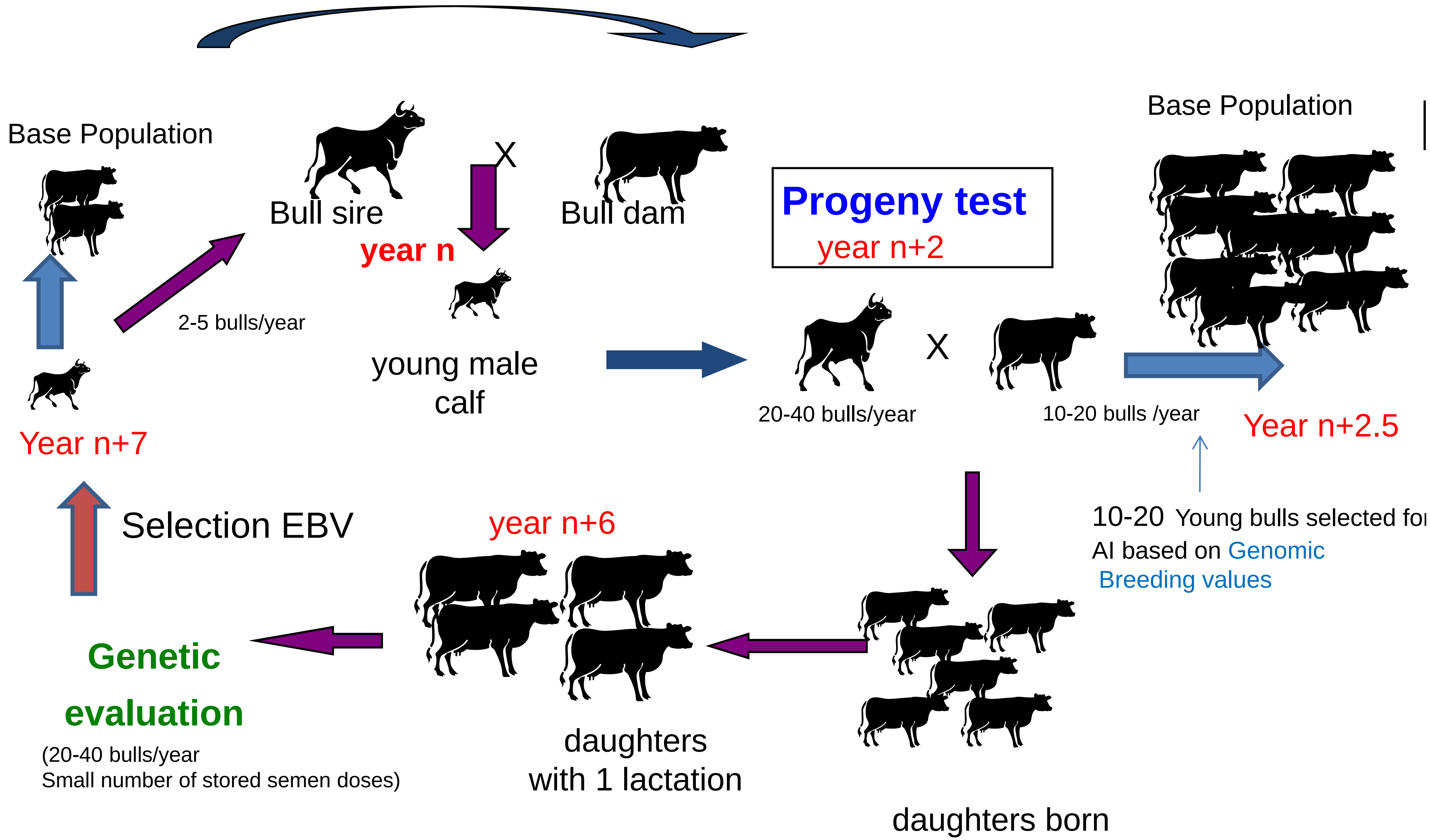
Disease Diag. Lab.

SMS – Alert Messages

Mobile

Farmer with animals in a village

Progeny testing scheme – Young Sire Model



Base Population

Bull sire

X

Bull dam

year n

2-5 bulls/year

young male calf

Year n+7

Progeny test

year n+2

20-40 bulls/year

X

10-20 bulls /year

Base Population

Year n+2.5

10-20 Young bulls selected fo AI based on Genomic Breeding values

Selection EBV

year n+6

Genetic evaluation

(20-40 bulls/year
Small number of stored semen doses)

daughters with 1 lactation

daughters born

Our challenges and way out

- Not having many progeny tested bulls with high reliability, our Reference population is too small.
- This means we cannot workout Genomic Breeding values based on the reference population of bulls.
- We propose to develop prediction equations based on small number of bulls whose breeding values we have and a large number of recorded female population.
- Recent studies have shown that the accuracy of Genomic breeding values of male calves could be raised up to 35-40% when genomic prediction equations are developed based on a reference population having small number of progeny tested bulls (around 100) and a large number of recorded females (around 2000).

Proposal for Genomic Selection

- It is proposed to develop predicton equations based on the available records of:
  - bulls (whose breeding values are available)
  - recorded females
- Two approches to estmate genomic breeding values would be tried:
  - (i) Direct estimation of marker effects using Bayes B approach
  - (ii) gBLUP using genomic relation matrix derived from markers among all individuals

What needs to be done to introduce Genomic Selection procedures

- Establish reference population of recorded male and female animals
- Collect biological samples of recorded animals
- Genotype / sequenced recorded populations – proposal is to set up a genomic service laboratory
- Standardize procedure for genomic breeding values estimation – proposal is to set up a bio-informatics unit
- Produce and select best bulls and bull mothers based on predicted genomic breeding values – proposal is to set up a nucleus herd of elite animals of various indigenous breeds having facilities for ET and IVF

Animals planned to be recorded in PT projects

Breed	**Project**	**Animals to be milk recorded under NDP projects**				
		14-15	15-16	16-17	17-18	Total
HF	KMF	2520	2537	2471	2885	10413
CB HF	SAG HF CB	2040	2082	2152	2467	8741
CB HF	ULDB	1800	1900	1900	1840	7440
CB HF	KLDB	2200	3000	3000	3000	11200
CB HF	BAIF	3200	3200	3200	3200	12800
CB Jersey	TCMPF	6400	8020	8052	8230	30702
CB Jersey	APLDA	1700	1660	1436	2091	6887
Murrah	SAG Murrah	1500	1940	1969	2269	7678
Murrah	PLDB	1700	1700	1700	1940	7040
Murrah	ABRO	1200	1200	1370	1607	5377
Murrah	HLDB	5000	5000	6000	6000	22000
Mehsana	Banas	1100	1100	1200	1200	4600
Mehsana	Mehsana	1100	1210	1311	1311	4932
	Total	**31460**	**34549**	**35761**	**38040**	**139810**

Animals planned to be recorded in PS projects

Breed	**Project**	**Animals to be milk recorded under NDP projects**				
		14-15	15-16	16-17	17-18	Total
Kankrej	Banas Union	700	400	300	300	1700
Gir	SAGB	500	400	300	300	1500
Jaffarabadi	SAGB	500	400	300	300	1500
Rathi	URMUL	700	600	600	600	2500
Hariana	HLDB	400	400	400	400	1600
Pandharpuri	MLDB	1000	700	600	600	2900
Nili-Ravi	PLDB	1000	800	700	500	3000
Tharparkar	RLDB	1000	800	700	500	3000
Sahiwal	PLDB/GANGMUL	1000	800	700	500	3000
Total		6800	5300	4600	4000	20700

Number of blood samples to be collected from female animals

Breed	14-15	15-16	16-17	17-18	18-19	Total
HF	2500	2500	2500	2000	2000	11500
HFCB	2000	2000	2000	2000	2000	10000
JCB	1000	2000	2000	2000	2000	9000
Indigenous cattle	2500	2500	2500	1500	1500	10500
Buffaloes	1000	2000	2000	2000	2000	9000
Total	9000	11000	11000	9500	9500	50000

Number of bulls whose semen doses to be stored

Breed	Year 1	Year 2	Year 3	Year 4	Year 5	Total
HF	40	40	40	50	50	220
HFCB	80	65	65	70	70	350
JCB	50	50	50	50	50	250
Indigenous cattle	40	40	40	30	30	180
Buffaloes	70	60	60	65	65	320
Total	280	255	255	265	265	1320

25 doses per bull will be stored

What additional genetic gain expected through Genomics selection procedures

- Maximum accuracy current PT project = 0.50
- Expected accuracy with genomic selection ~ 0.60 -0.70
- This means, when whole genomic selection procedures are used, about 20 to 40% increase in accuracy is expected over pedigree selection.
- In other words about 20-40% increase in genetic gain is expected over current progeny testing programmes.

Thank You