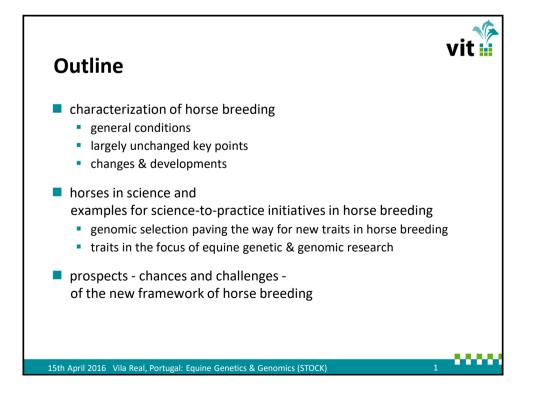IT-Solutions for
Animal Production

**4th International Meeting on Equine Production
on April 14-16, 2016, in Vila Real / Portugal**
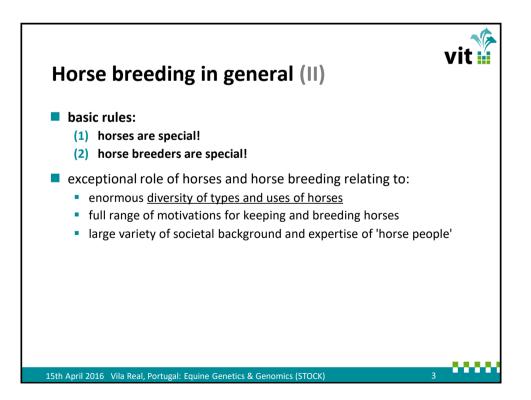
Equine Genetics & Genomics
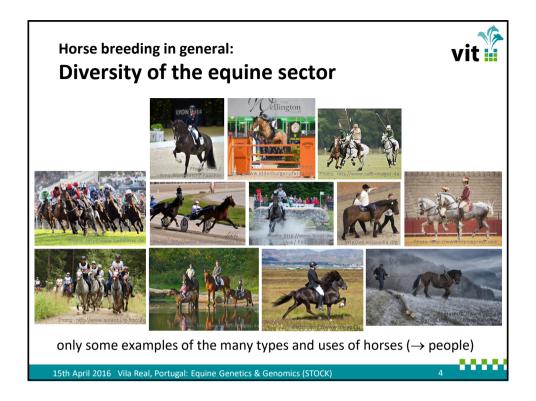
Challenges and chances of horse breeding
relating to new tools and genomic applications

K.F. Stock

IT Solutions for Animal Production (vit), Verden, Germany

Outline

- characterization of horse breeding
  - general conditions
  - largely unchanged key points
  - changes & developments

- horses in science and
  examples for science-to-practice initiatives in horse breeding
  - genomic selection paving the way for new traits in horse breeding
  - traits in the focus of equine genetic & genomic research

- prospects - chances and challenges -
  of the new framework of horse breeding

Horse breeding in general (I)

- **basic rules:**
  - **(1)** **horses are special!**
  - **(2)** **horse breeders are special!**

BE CAREFUL WITH:
- ➢ comparing horses with 'classical' livestock species (like dairy cattle)
- ➢ comparing horse keeping and breeding with respective pet sectors
- ➢ comparing horse breeders with other animal breeders
- ➢ discussing horse breeding with animal breeders without horse affinity

Horse breeding in general (II)

- **basic rules:**
  - **(1)** **horses are special!**
  - **(2)** **horse breeders are special!**
- exceptional role of horses and horse breeding relating to:
  - enormous diversity of types and uses of horses
  - full range of motivations for keeping and breeding horses
  - large variety of societal background and expertise of 'horse people'

Horse breeding in general:
Diversity of the equine sector

vit

only some examples of the many types and uses of horses (→ people)

Horse breeding in general (III)

vit

- **basic rules:**
    - **(1) horses are special!**
    - **(2) horse breeders are special!**

- exceptional role of horses and horse breeding relating to:
    - enormous <u>diversity of types and uses of horses</u>
    - full range of motivations for keeping and breeding horses
    - large variety of societal background and expertise of 'horse people'

- long tradition and cultural value of horse breeding
    - stabilizing factor
    - stimulus of continued support and engagement
    - conflict potential

Key points in horse breeding (I)

- occupying and intense in many respect
  - time, space / surroundings, man-power, knowledge / expertise, money
  - long gestation length, long raising period,
    long training up to advanced level performance
  - daily management requirements, daily exercise requirements, ...

> ➢ horse breeders as ambitious people
> willing to invest (a lot) to gain (a lot) and
> with often emotional rather than purely rational decision making

Key points in horse breeding (II)

- occupying and intense in many respect
  - time, space / surroundings, man-power, knowledge / expertise, money
  - long gestation length, long raising period,
    long training up to advanced level performance
  - daily management requirements, daily exercise requirements, ...

> ➢ horse breeders as ambitious people
> willing to invest (a lot) to gain (a lot) and
> with often emotional rather than purely rational decision making

- diverse and usually complex breeding goals
  - conformation
  - health
  - behavior
  - performance

Changes & developments

- increased internationality of horse breeding
  - exchange of information
  - global rather than local or national comparisons
  - exchange of genetic material (horses, semen)

➤ increased competition among breeders and studbooks

- structural changes of society also affecting the equine sector
  - economic development
  - changed background of horse breeders (familial heritage ↓, agriculture ↓, sport / leisure riding ↑, hobby horse keeping ↑)

➤ new expectations and requirements
  in times of increased pressure and limited resources

➤ high motivation to re-consider and possibly initiate changes

---

Recent survey: Sport horse studbooks (I)

Table 3
Traits included in the individual verbal breeding objective definitions presented by 19 breeding organisations in Europe

	Conformation (n=17)	Show jumping (n=16)	Dressage (n=14)	Gaits (n=14)	Behaviour (n=11)	Eventing (n=11)	Health[a] (n=9)	Driving (n=4)	Fertility (n=3)
BAD[b]	X	X	X						
BAVAR	X	X	X						
BWP	X								
DWB	X	X	X						
FWB	X	X	X						
HAN	X	X	X						
HOLST	X	X							
HUN	X	X	X						
ISH	X	X	X						
KWPN	X	X	X						
NRPS	Unspecific								
NWB	X	X	X						
OLD	X	X	X						
SF	X	X	X						
SHBGB	X	X	X						
SI	X								
SWB	X	X	X						
TRAK	X								
WEST	X	X	X						

[a] Including durability and soundness.

Source: Koenen et al. 2004

Trait	N	Mean score (scale 0-3)	Score counts irrelevant (0)	Score counts highly important (3)
Conformation	14	2.286	0	5
Gaits	14	2.286	1	8
Jumping ability	14	2.571	1	11
Dressage	14	2.214	1	8
Show-jumping	14	2.571	1	11
Eventing	13	1.154	4	3
Driving	13	0	13	0
Allrounder qualities	12	1.417	2	1
Behavior and temperament	13	2.615	0	8
Health / soundness and durability	13	2.385	0	7
Fertility / reproductive performance	13	1.769	2	4

Source: Stock et al. 2015

- importance of traits / trait groups:
  - relevance of conformation
  - strength of health and behavior aspects

Slide 10

Recent survey: Sport horse studbooks (II)

vit

Source: Stock et al. 2015

- importance of traits / trait groups
- development of focuses in breeding:
  - highest expectations with regard to relevance of health traits

Slide 11

Equine health in science & practice (I)

vit

- health and durability as integral parts of breeding programs (today as in the past)

- discrepancy between importance of health aspects (for all parties) and their consideration in breeding
  - limited documentation and regulation of inherited defects and diseases in breeding animals (NIKOLIC 2009)
  - mostly inefficient selection strategies (if any):
    - mainly phenotype-based selection (mass selection), i.e. exclusion of affected individuals
    - often only in stallions, but not in mares

- promising results of broader pilot studies and in-depth research
  - X-ray screening data (radiographic health of the limbs)
  - clinical examination results

Equine health in science & practice (II)

Why is there this obvious discrepancy?

- reasons for little progress in breeding for improved health
  - sensitivity of data, data security and data privacy
  - necessary integration of non-breeding (= veterinary medical) data sources in the breeding routines
  - expected very slow progress due to low heritabilities implying need for lots of data to get reliable support of breeding decision

➢ lack of data due to difficult and expensive health data collection

Equine health in science & practice (III)

Why is there this obvious discrepancy?

➢ lack of data due to difficult and expensive health data collection

Why do we observe signs of optimism now?

- reasons for (new) engagement for better health-oriented selection
  - positive experiences with integrated, interdisciplinary data base system (better = more powerful and safer technology)
  - new generation of breeders, owners, buyers, ... with increased expectations regarding sustainable breeding programs
  - new possibilities to develop and run efficient health improvement programs thanks to the <u>new genomic tools</u>

'New era of breeding'

- **dilemma of horse breeding:** most interesting traits belonging to the most <u>challenging traits</u> for animal breeders
  - low heritability
    non-genetic influences complicating and delaying breeding progress
    (what you see = phenotype only vaguely reflects the underlying genetics)
  - difficult and expensive data collection
    rareness of high quality phenotype data
  - mostly old age of animals when getting reliable genetic proofs
  - e.g. health traits, behavior aspects, advanced level performance traits
- **genomic selection** 'opening the window' for the challenging traits
  - possible focusing of phenotypic data collection
  - very early accessibility of reliable predictors of the genetic potential
    ($\rightarrow$ shortening of the generation interval)

15th April 2016   Vila Real, Portugal: Equine Genetics & Genomics (STOCK)   14

Information basis for selection

Results of model calculations:
accuracy of dressage EBVs as function of age of an elite stallion
(EBV = estimated breeding value)   MARK 2013

Accuracy (%)

EBV w. 5000 genotypes

EBV w. 500 genotypes

first offspring start competing
(progeny performance records)

Traditional EBV

stallion starts competing
(own performance records)

Stallion age (years)

- genomic selection implying significant gain in time and accuracy
  $\rightarrow$ breeding progress $\uparrow$

15th April 2016   Vila Real, Portugal: Equine Genetics & Genomics (STOCK)   15

From genetics to genomics

- support of breeding by direct reference to genomic data
  - reading the genetic code and comparing it between individuals
  - <u>not</u> manipulating the genetic code, but applying statistical methods to interpret specific parts of it

- (relatively) old idea of searching for what makes 'the genetic potential' of an individual

- routine genomic applications becoming available only recently
  - human molecular genetic research as driver
  - enormous technological progress, followed by dropping costs, expansion of logistics, development of testing systems for routine use, development of statistical methodology, ...

What is genomic data?

- genome
  - blue-print of life processes
  - arranged in chromosomes

- genetic code
  - inscribed in DNA (double helix with 'ladder structure')
  - genetic alphabet: G, C, A, T
  - genome size (mammals): about 2.5-3.0 x 10^9 base pairs
  - coding sequence (genes): ca. 1 gene per 100,000 bases

Slide 18:

Rapid growth of genomic research

- genome sequence information as precondition of systematic working with genomic data

Cost per Genome

Time scale of full genome sequence completion:

2001 — 2003 dog — 2004 chicken — 2007 horse — 2008 sheep — 2009 cattle pig rabbit turkey — 2010 — 2011 (atlantic salmon) — 2012 goat

15th April 2016   Vila Real, Portugal: Equine Genetics & Genomics (STOCK)   18

Slide 19:

How to use genomic data (I)

- variation in the genetic code (genomic variation)
  - 'game field' of evolution and utilizable by breeders
  - different types and effects
  - possible use as genetic markers
- SNP
  - most abundant variant type → whole genome coverage
  - dense marker maps, libraries → sets of informative SNPs
  - mostly biallelic (base 1 or 2) → easy to assay and interpret
  - SNPs as 'stars of the genomic era'!

SNP 1   SNP 2   SNP 3

```
1: acggttagctacaattatttaaacggggaggagggattttattaaccagatgtg
2: acggttatctacaattatttaaacgggaggagggattttattaaccagatgtg
3: acggttaactacaattatttaaacgggaggagggattttattaaccagatgtg
4: acggttaactacaattatttaaatgggaggagggattttattaaccaaatgtg
5: acggttatctacaattatttaaatgggaggagggattttattaaccaaatgtg
6: acggttatctacaattatttaaatgggaggagggattttattaaccaaatgtg
```

Examples of SNP reads of 6 animals (SNP = single nucleotide polymorphism)

SNP array or SNP chip
(for the horse e.g. Equine 70k SNP chip)

15th April 2016   Vila Real, Portugal: Equine Genetics & Genomics (STOCK)   19

How to use genomic data (II)

gene | genetic marker

- idea behind using genomic data via genetic markers:
  if you <u>screen</u> enough sites spread across <u>the whole genome</u> (SNPs!),
  some will be close enough to genes (relevant for traits of interest)
  to show similar distribution patterns

How to set-up genomic selection? (I)

ID	SNP genotype	phenotype	dGV
animal A	...AGGCACC GCAATCCACG GAGGCTACGC CCTCACCGGA GGTTTCGCTC TCCACGG... ...AGGCACC GCAATCCACG GAGGCTACGC CCTCACCGGA GGTTTCGCTC TCCACGG...	-	--
animal B	...AGGCACC GCAATCCACG GAGGCAACGC CCTCACCGGA GGTTTCGCTC TCCACGG... ...AGGCACC GCAATCCACG GAGGCAACGC CCTCACCGGA GGTTTCGCTC TCCACGG...	+++	++
animal C	...AGGCACC GCAATCCACG GAGGCTACGC CCTCACCGGA GGTTTCGCTC TCCACGG... ...AGGCACC GCAATCCACG GAGGCAACGC CCTCACCGGA GGTTTCGCTC TCCACGG...	+	∅
animal D	...AGGCACC GCAATCCACG GAGGCAACGC CCTCACCGGA GGTTTCGCTC TCCACGG... ...AGGCACC GCAATCCACG GAGGCAACGC CCTCACCGGA GGTTTCGCTC TCCACGG...	++	++

favorable SNP allele (SNP effect > 0): A
favorable SNP allele (SNP effect < 0): T

accordingly for N SNP

summation of SNP effects

Phenotype example:
Body size (withers height)
from short (-) to very tall (+++)

- **step I:** linking phenotypic and genomic data (<u>reference population</u>)
  → analyzing distribution pattern → deriving prediction formula
- **step II:** using genomic data (SNP genotypes) only for prediction

How to set-up genomic selection? (II)

ID	SNP genotype	phenotype	dGV
animal E	...AGGCACC GCAATCCACG GAGGCTACGC CCTCACCGGA GGTTTCGCTC TCCACGG... ...AGGCACC GCAATCCACG GAGGCTACGC CCTCACCGGA GGTTTCGCTC TCCACGG...	?	-
animal F	...AGGCACC GCAATCCACG GAGGCAACGC CCTCACCGGA GGTTTCGCTC TCCACGG... ...AGGCACC GCAATCCACG GAGGCAACGC CCTCACCGGA GGTTTCGCTC TCCACGG...	?	+++
animal G	...AGGCACC GCAATCCACG GAGGCTACGC CCTCACCGGA GGTTTCGCTC TCCACGG... ...AGGCACC GCAATCCACG GAGGCTACGC CCTCACCGGA GGTTTCGCTC TCCACGG...	?	+
animal H	...AGGCACC GCAATCCACG GAGGCAACGC CCTCACCGGA GGTTTCGCTC TCCACGG... ...AGGCACC GCAATCCACG GAGGCAACGC CCTCACCGGA GGTTTCGCTC TCCACGG...	?	++

Prediction of genetic background of the phenotype

SNP effect according to genotyping results
(favorable / unfavorable alleles)

accordingly for N SNP

summation of SNP effects

- **step I:** linking phenotypic and genomic data (reference population)
  → analyzing distribution pattern → deriving prediction formula
- **step II:** using genomic data (SNP genotypes) only for prediction

---

Genetics vs. genomics:
Improved individual characterization

SNPs

Illustration of individual genomic differences between full sibs
(tracing chromosomes through generations from left to right:
offspring - parents - grand parents - great grand parents)

BV_dam = 110   BV_sire = 134

Possible egg-cells   Possible sperm-cells

Average = 55 = ½ × BV_dam   Average = 67 = ½ × BV_sire

Population genetic approach of working with averages
(until individual information becomes available)

Source: STOCK ET AL. 2016

Genomic selection in practice?

- gene tests (direct, indirect) as standard laboratory service
  - colors
  - genetic defects, diseases
  - performance aspects, e.g. gaiting ability and trotting race performance
- first offers of genomic evaluations
  - racing performance / Thoroughbred
  - osteochondrosis / Warmblood (?)
- several projects initiatives to:
  a. improve the phenotype data basis
     - health
     - conformation and performance
  b. implement genomic selection
     - national (e.g. GenHors / DK)
     - international (consortium formation underway / sport horse)

EXAMPLE: Body size (I)

- model trait for studying genetic influences on quantitative traits (within and across species)
- importance across breeds of horses
  - distinct types of equine skeletal morphology
  - specific body size and shape variation patterns within and across breeds
- genome-wide association studies (GWAS)

EXAMPLE: Body size (II)

Figure: Manhattan plot for height at withers based on dEBV. (Signer-Hasler et al., 2012)

Figure: Estimates of height at withers dEBV variance explained by SNPs on 31 autosomes (black bars) and the two identified QTL (green bars). (Signer-Hasler et al., 2012)

- 1.077 **Franches-Montagnes** horses (212 stallions, 865 mares):
  height at withers (WH) of 145-165cm, de-regressed estimated breeding values for WH (dEBV)
- Illumina EquineSNP50® BeadChip (54,602 → 38,124 SNPs)
- QTL regions on ECA3 (*LCORL/NCAPG*) and ECA9 (*ZFAT*) - together explaining 18.2 % of dEBV variance

EXAMPLE: Body size (III)

- recent increase of knowledge on genetic size regulation:
  LCORL (ligand-dependent nuclear receptor compressor-like protein)

  - consistent GWAS results in horses
    Franches-Montagne horses (Signer-Hasler et al. 2012)
    German warmblood horses
      (Kühn et al. 2012 / Tetens et al. 2013; Metzger et al. 2013a,b)
    Arabs (Ricard et al. 2013)
    multiple breeds and breed types (Metzger et al. 2013b)

  - relevant size regulator across species
    humans (Lango Allen et al. 2010), cattle (Pryce et al. 2011)

Figure: Manhattan plot for height at withers based on measurements. (Tetens et al., 2013)

EXAMPLE: Locomotion (I)

- highly complex interplay of peripheral and central nerval function
  - some very specific functional defects
  - **understanding of regular function** (across species) often driven by studies of **specific function** (in target and model species)

gaitedness and high-speed trotting

- breed-specific ability to perform alternate gaits, i.e. surplus to basic gaits (walk, trot, canter): pace, ambling gaits (regular rhythm, lateral, diagonal ambling)
- breed-specific ability to sustain trot at high speed, i.e. delay transition into canter
- related to general control / coordination of limb movement

Photo: Andersson, 2012

EXAMPLE: Locomotion (II)

A Mutation in a Novel Transcription Factor Affects the Pattern of Locomotion in Horses

- 70 **Icelandic horses** (40 with pace/5-gaited, 30 without pace/4-gaited); **capability of pace** (Y/N, competition records)
- Illumina EquineSNP50® BeadChip (54,602 SNPs)
  → GWAS: strong association signal on ECA23

Table: Distribution of SNP genotypes. (Andersson et al., 2012)

	CC/CT	TT
pace	1	39
not pace	21	9

Figure: GWAS results for the ability to pace with genome-wide significant SNP on ECA23. (Andersson et al., 2012)

EXAMPLE: Locomotion (III)

A Mutation in a Novel Transcription Factor Affects the Pattern of Locomotion in Horses

nature

- 70 **Icelandic horses** (40 with pace/5-gaited, 30 without pace/4-gaited); **capability of pace** (Y/N, competition records)
- Illumina EquineSNP50® BeadChip (54,602 SNPs)
  → GWAS: strong association signal on ECA23
- for refinement: 2 **Icelandic horses**, 1 with pace/5-gaited (AA), 1 without pace/4-gaited (CC)
- sequence data → mutation detection: nonsense mutation C/A (transcription factor DMRT3)
  → premature stop codon (truncated protein, 174 amino acids less)
- mouse model (wild-type and Dmrt3-null mice)
- DMRT3 gene expression: subset of inhibitory neurons in spinal cord (crossing dorsal midline, connecting to motor neurons)

Characterization of motor coordination in mice lacking Dmrt3 ANDERSSON ET AL. 2012

EXAMPLE: Locomotion (IV)

A Mutation in a Novel Transcription Factor Affects the Pattern of Locomotion in Horses

nature

- ...
- 352 **Icelandic horses**, 176 **gaited horses** (6 breeds), 218 **non-gaited horses** (8 breeds), 414 **horses bred for harness racing**
- high frequencies of DMRT3 mutation in gaited and harness racing horses
- DMRT3 mutation: permissive for performing alternate gaits, i.e. pace or four-beat ambling gaits, and high-speed trot ('gait keeper mutation')

Table: Allele frequency of the DMRT3 nonsense mutation among horse populations. (Andersson et al., 2012)

Breed	n	p(A)
Icelandic horses*		
Four-gaited†	124	0.65
Five-gaited	66	0.99
Random sample	162	0.89
Other gaited horses		
Kentucky mountain saddle horse	22	0.95
Missouri fox trotter	40	1.00
Paso fino	45	1.00
Peruvian paso	19	1.00
Rocky mountain horse	17	1.00
Tennessee walking horse	33	0.98
Non-gaited horses		
Arabian horse	18	0.00
Gotland pony	28	0.00
North-Swedish draft horse	31	0.00
Przewalski's horse	6	0.00
Shetland pony	20	0.00
Swedish ardennes	22	0.00
Swedish warmblood	64	0.00
Thoroughbred	29	0.00
Horses bred for harness racing		
Standardbred, trotter (Sweden)	270	0.97
Standardbred, trotter (USA)	57	1.00
Standardbred, pacer (USA)	40	1.00
French trotter (France)	47	0.77

Table: Distribution of the DMRT3 mutation in the Icelandic horse. (Andersson et al., 2012)

Phenotype	CC	CA	AA	Total
Five-gaited	0	1	65	66
Four-gaited	2	83	39	124
Total	3	105	149	257

$p = 2.4 \times 10^{-14}$

EXAMPLE: Locomotion (V)

Figure: Frequencies of alleles A (red line) and C (green line) in DMRT3 in the Icelandic horse population 1980-2012 (Kristjansson et al. 2014)

- 667 **Icelandic horses**: gaitedness, gait quality, speed
- DMRT3 genotyping
- positive selection for DMRT3 mutation (selection advantage): A allele frequency ↑ (in birth year 2012: 94%), expected elimination of wildtype allele C

EXAMPLE: Locomotion (VI)

Figure: Frequency distribution of the DMRT3 gait-altering mutation for breeds with at least five samples tested (Promerová et al. 2014)

- 4,396 **horses from 141 breeds**
- DMRT3 genotyping
- detection of DMRT3 mutation in 68 breeds with frequencies of 1-100%, high frequencies (>50%) in gaited breeds and trotters

EXAMPLE: Racing performance (III)

Prospects of horse breeding

- improvement of routine breeding applications implying:
    - increased need for high quality phenotype data
    - new tools → new traits
        - pushing health data collection
        - replacement of subjective scores by more objective linear traits
    - opportunity to optimize the trait spectrum in the breeding program (challenging breeding goal traits ↑)
    - needs for improved support and advisory service for breeders (genomic selection is a 'sharp knife'!)
- high motivation to intensify collaboration
    - synergistic effects
    - long-term benefit of improved conditions for implementing **new genomic tools with enormous potential for the equine sector**

> *'In the era of genotype, phenotype is king'*
> COFFEY 2010