Joint estimation of additive and dominance effects of markers using a genomic model with a residual polygenic effect

Zengting Liu, H. Alkhoder, F. Reinhardt and R. Reents
vit, Germany, http://vit.de

Introduction

- Routine genomic evaluation for German Holsteins
  - Bull reference population with high reliable daughter proven bulls
  - Indirect phenotype: deregressed EBV of conventionally evaluated traits
  - No cows included in DEU Holstein reference population

- Advantages of cow reference population for genomic evaluation
  - Avoid impact of genomic pre-selection on male side (Schaeffer, 2014)
  - Bull reference population available very late for novel traits

- New opportunities by using cow reference population
  - Explore non-additive genetic effects (Toro & Varona, 2012)
  - Additionally estimate dominance effects (Su et al. 2012)
  - Use dominance effects for genomic mating (Sun et al. 2013)
  - Mate allocation with good combining ability in commercial herds

Introduction

- Dominance effects in conventional genetic evaluation
  - Usually ignored, also for low heritable traits
  - Large full-sib families required for accurate dominance estimates
  - Pedigree information not informative enough
  - Greater computational complexity

- Two alternative genomic models with dominance effects
  - Su et al. (2012)'s model with pure additive genetic effects
  - Vitezica et al. (2013)'s model with substitution/breeding value effects

- Inclusion of a residual polygenic effect for reducing the inflation of genomic prediction (Liu et al. 2011)

- Objectives: to develop and test a software for a genomic dominance model with a residual polygenic effect
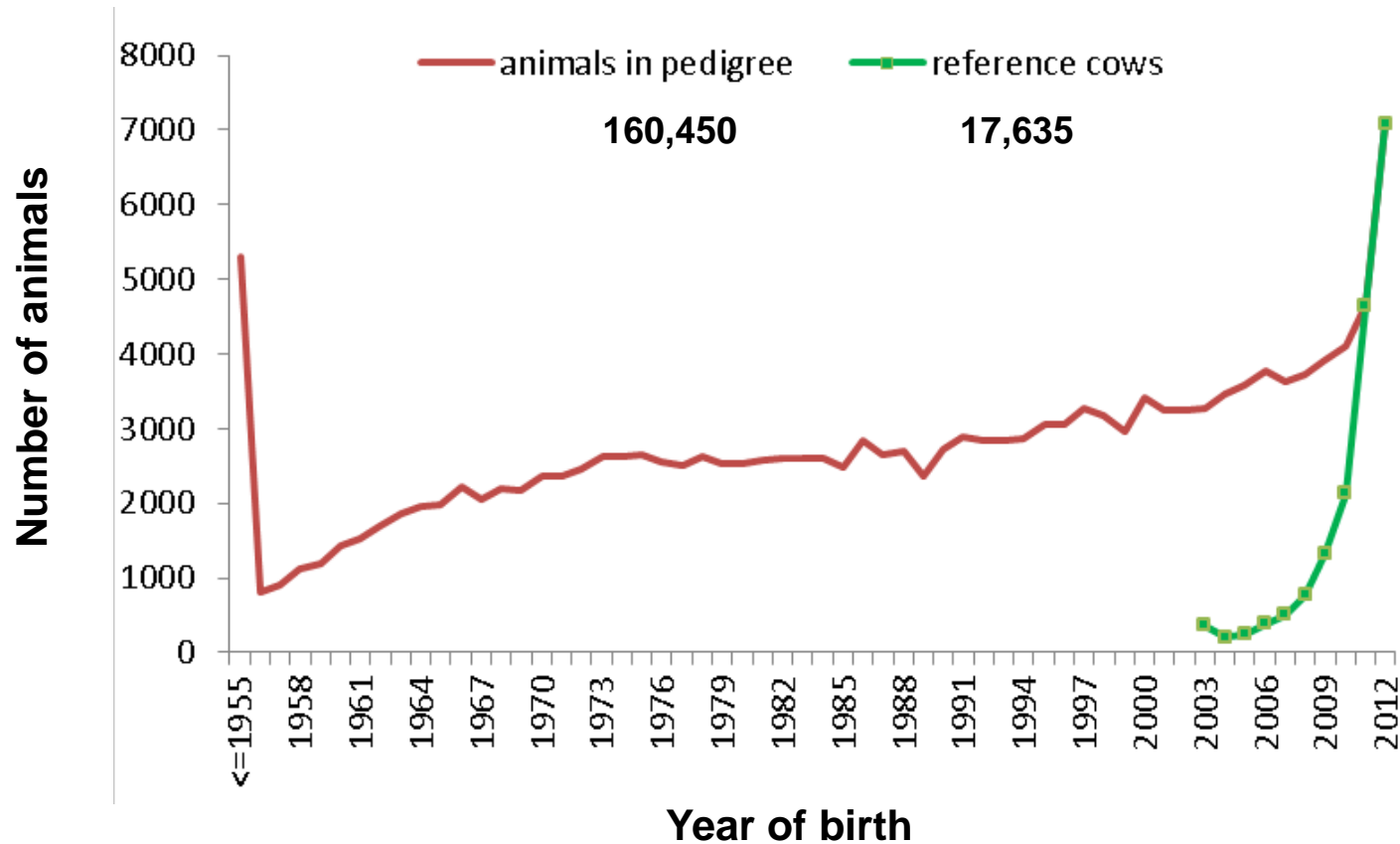
Data materials

- Phenotypes: cow lactation yield deviations of a test-day model (Liu et al. 2004) as TEST TRAITS
  - Traits: first lactation milk, fat and protein yields and SCS

- In contrast to 'indirect phenotypes' deregressed proofs, yield deviations can contain non-additive genetic effects, if exist

- All available genotyped German Holstein females by April 2015
  - Including selectively genotyped elite cows or bull dams
  - Original 50K or EuroG10K genotypes imputed to 50K

- Some old cows were not suited for a cow reference population: selectively genotyped females
  - Nonetheless acceptable for testing the model and software

Cow reference population and pedigree animals

Joint estimation of substitution and dominance effects together with residual polygenic effects

- $$y_i = \mu + u_i + \sum_{j=1}^{m} z_{ij}\alpha_j + \sum_{j=1}^{m} w_{ij}d_j + e_i$$

  y_i is yield deviation of cow with EDC φ_i

  u_i is residual polygenic variance $\text{var}(u_i) = k\sigma_g^2$

  $$\text{var}(e_i) = \sigma_e^2 / \varphi_i$$

- $z_{ij} = \dfrac{n_A - 2p_j}{\sqrt{2p_j(1-p_j)}}$ is regression on substitution effects of SNP markers α_j

  (n_A = 2 / 1 / 0 for AA / AB / BB)

- Regression on dominance effect d_j (*Vitezica et al. 2013*)

  For AA genotype: $w_{ij} = \dfrac{-2(1-p_j)^2}{2p_j(1-p_j)} = -\dfrac{1-p_j}{p_j}$

  For AB genotype: $w_{ij} = \dfrac{2p_j(1-p_j)}{2p_j(1-p_j)} = 1$

  For BB genotype: $w_{ij} = \dfrac{-2(1-p_j)^2}{2p_j(1-p_j)} = -\dfrac{p_j}{1-p_j}$

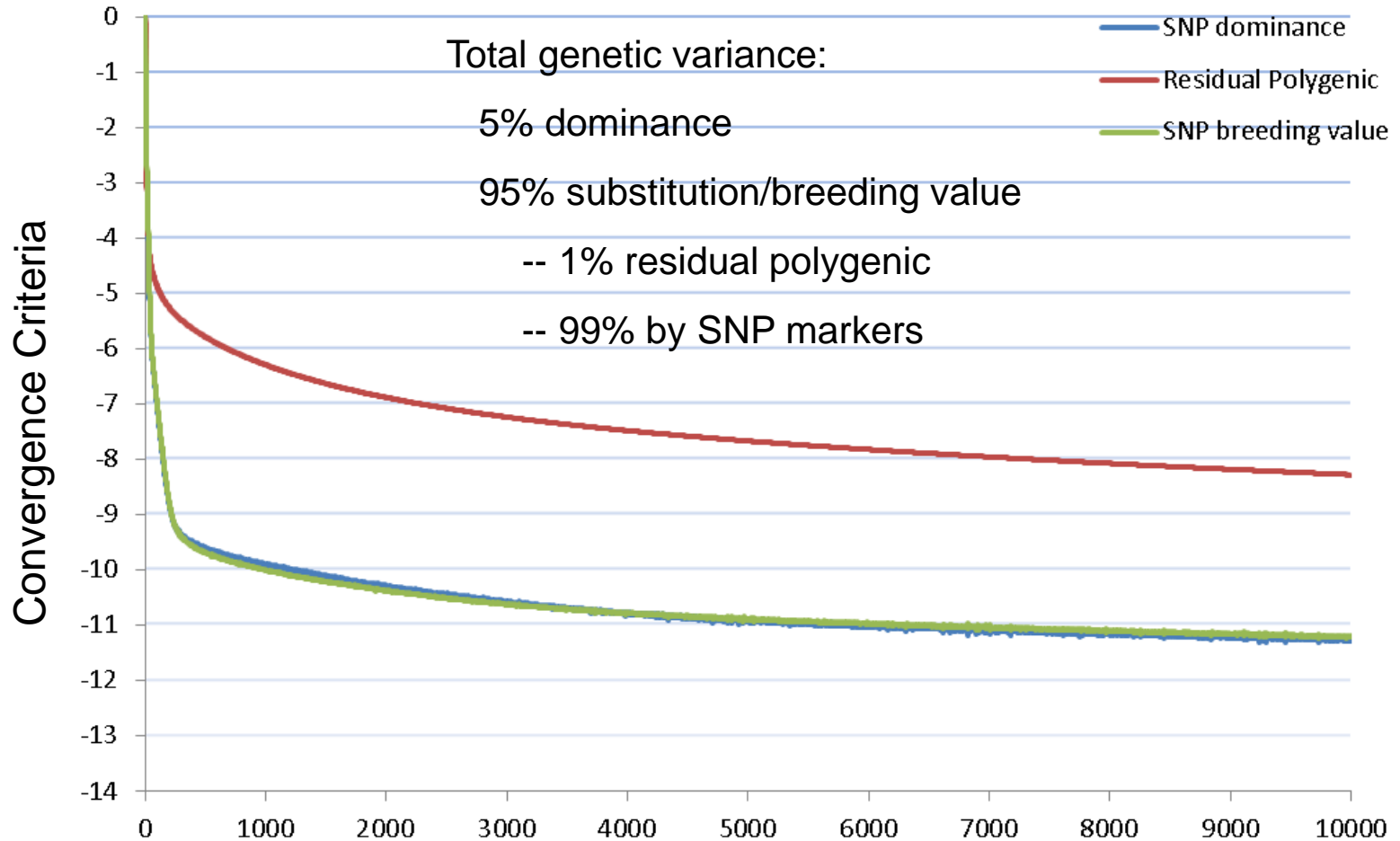# Two alternative genomic dominance models

- Two genomic models for estimating dominance effects
  - 'Biological genotypic' additive effects of markers by Su et al. (2012)
  - 'Statistical' breeding values of markers by Vitezica et al. (2013)
  - Both models are equivalent and convertible
  - For consistency with conventional definition of (residual) polygenic effect, Vitezica et al. model was preferred

- Two scenarios with regard to dominance variance
  - 5% and 10% of total genetic variance were assumed

- CPU and RAM usage
  - Linux server with Intel Xeon CPU E5-2690 v2 @ 3.00GHz
    - 20 cores and 512 Gb RAM
  - ~ 6.5 Gb RAM, mainly for storing the regression coefficients
  - ~ 11 hours x 20 cores for 10,000 rounds

Results: Rate of convergence

Total genetic variance:

5% dominance

95% substitution/breeding value

-- 1% residual polygenic

-- 99% by SNP markers

Trait: first lactation milk yield   17,635 reference cows, 160,250 animals in pedigree
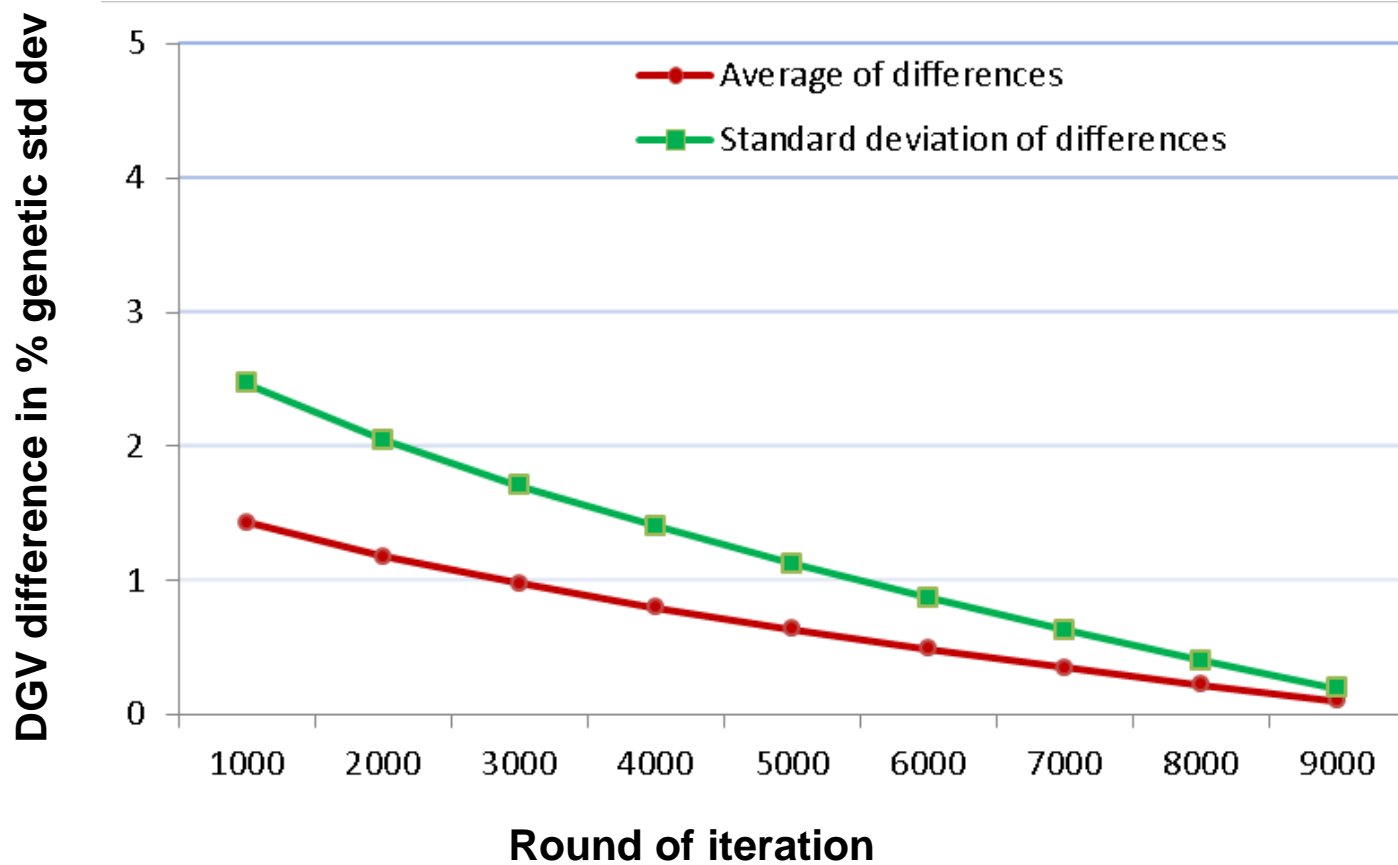
Convergence criteria: $\log(\Sigma(Snew-Sold)^2/ \Sigma Snew^2)$

Results: Breeding value SNP and DGV effects

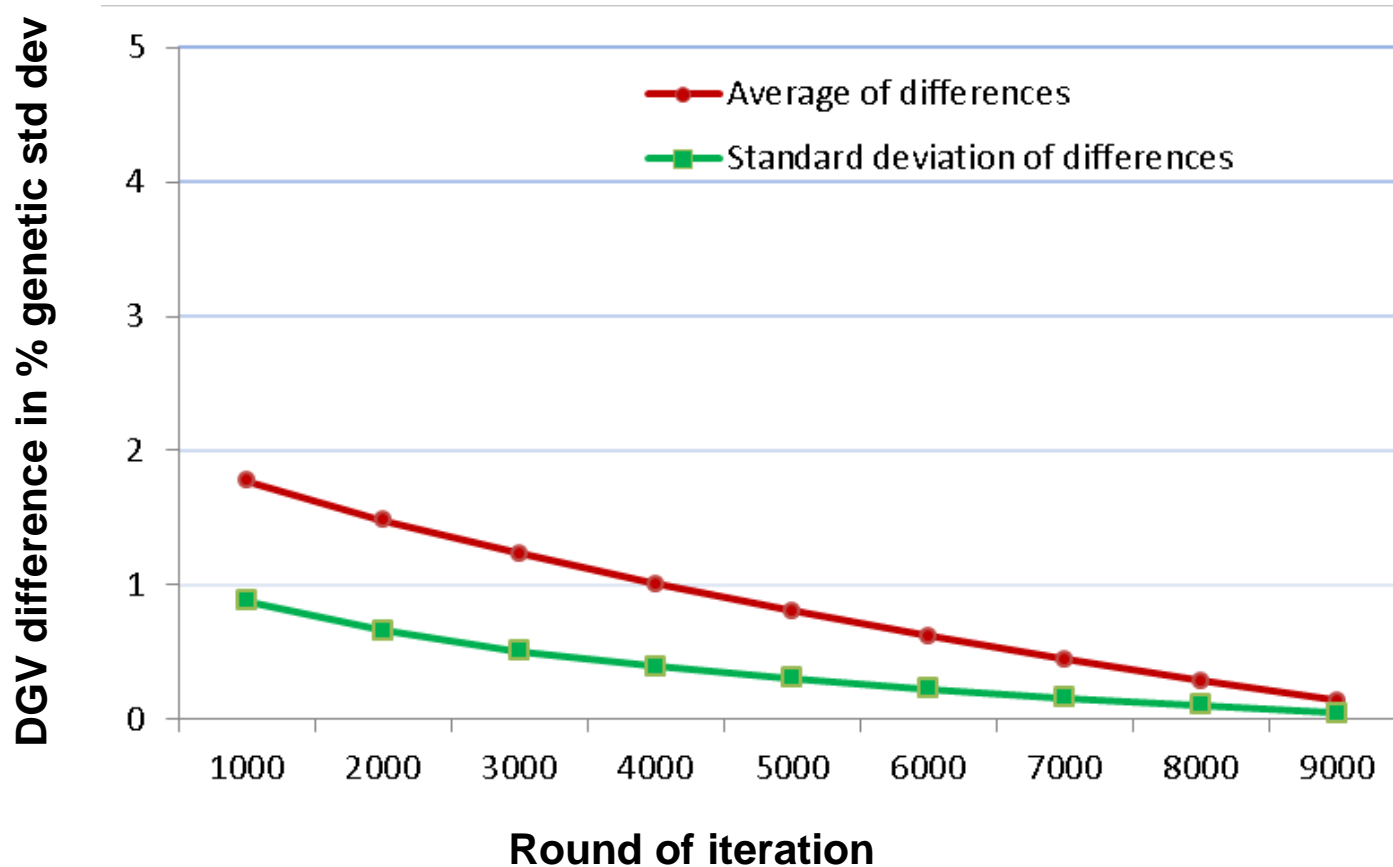Increasing SNP effect or DGV correlations with final round, 0.99984 at round 1000
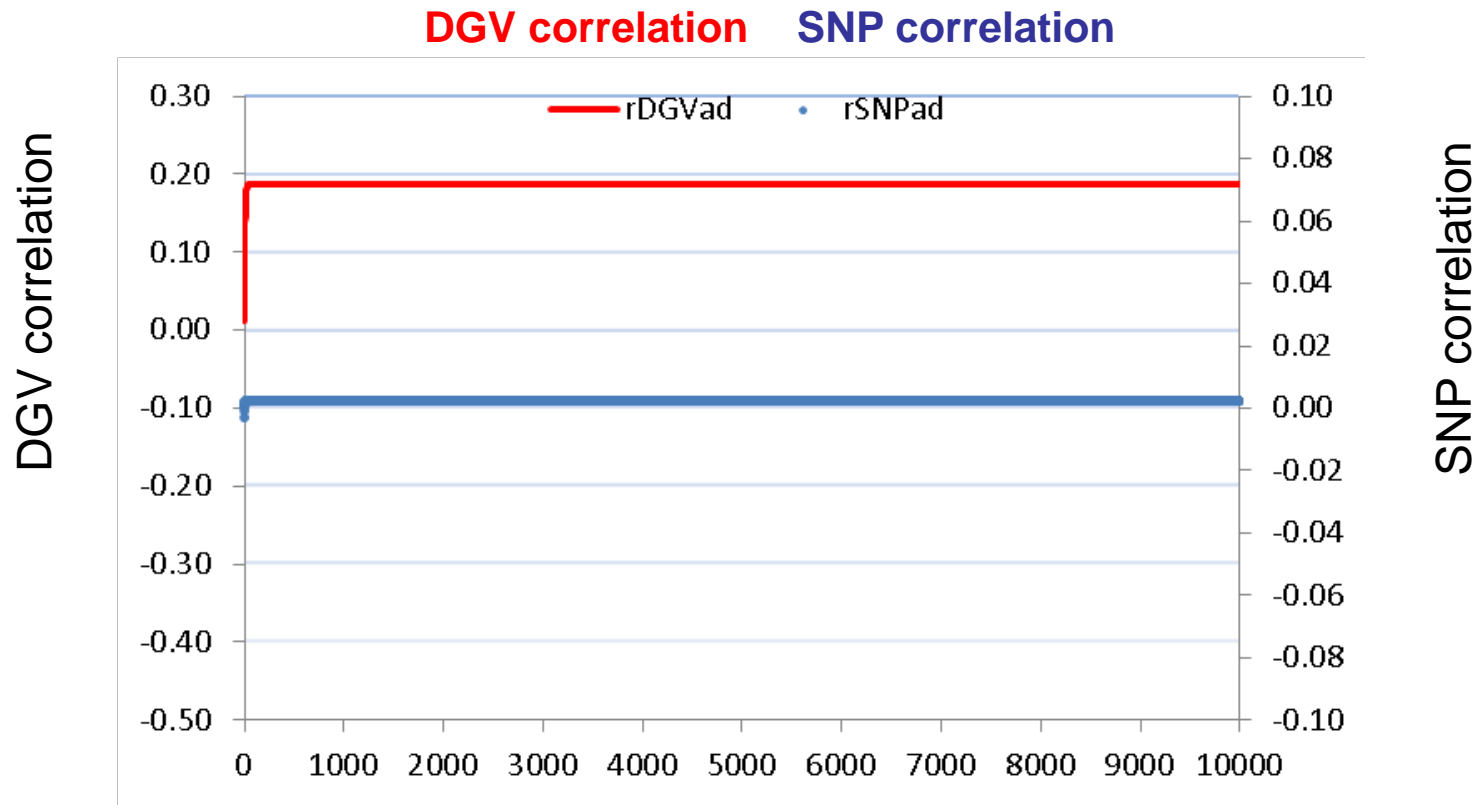
Differences to the final round 10,000

Results: Dominance SNP and DGV effects

Increasing SNP effect or DGV correlations with final round, 0.99986 at round 1000

Differences to the final round 10,000

Results: Correlations of breeding value with dominance effects on SNP and DGV levels
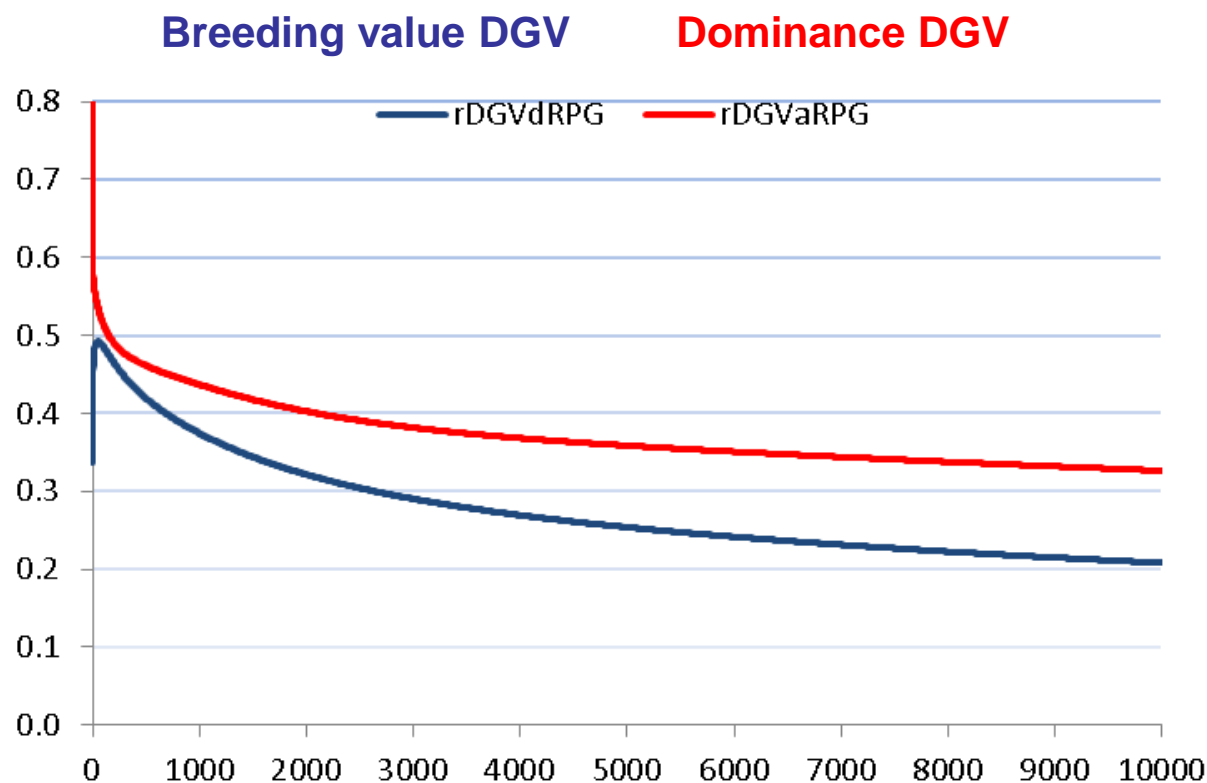

Trait: first lactation milk yield   17,635 reference cows, 160,250 animals in pedigree

5% dominance variance

Results: Correlations of DGVs with residual polygenic effects

Breeding value DGV   **Dominance DGV**

Trait: first lactation milk yield   17,635 reference cows, 160,250 animals in pedigree

5% dominance variance

Results: the two genomic dominance models

- **Vitezica** et al (2013) with marker breeding values

- **Su** et al (2012) with pure additive effects of makers

- Little difference in rate of convergence between the two models

- Correlations of SNP effects: dominance 0.918, additive/breeding value 0.987

- Correlations of DGV: dominance 0.849, additive/breeding value 0.980, sum=dominance + additive/breeding value 0.998

Summary and conclusions

- In conventional evaluation, dominance effects were difficult to be accurately estimated based on pedigree info (Vitezica et al. 2013)

- However, using genomic data
    - Both effects of SNP markers were accurately estimated at round 1000
    - Easy to separate both effects, as correlations between breeding value and dominance effects quickly stabilized on both SNP and DGV levels

- More iteration rounds needed due to residual polygenic effects
    - Correlations stabilized rather slowly
    - As a result of very deep pedigree structure
    - But the overall impact is limited due to its low variance

- Estimating additive and dominance effects of SNP markers can be easily done much faster on multiple cores

Summary and conclusions

- **Vitezica** et al's breeding value model (2013) is preferred, as it is consistent with classical definition in conventional evaluation

- Extension to other traits, specially those novel with low heritability

- Predictive ability of the dominance model to be validated
  - Comparison to the current additive-effect only SNP model

- Genomic mating with dominance effects easier to do than in case of mating with conventional evaluation (Sun et al. 2013)

- Optimal dominance and residual polygenic variances for cow reference population to be determined via genomic validation

Acknowledgements

- German R&D project Kuh-L for financing

- Dr. Guosheng Su for discussing about the dominance models