Berner Fachhochschule
Haute école spécialisée bernoise
Bern University of Applied Sciences

QUALITAS

Genomic prediction in cattle based on sequence data

Mirjam Frischknecht

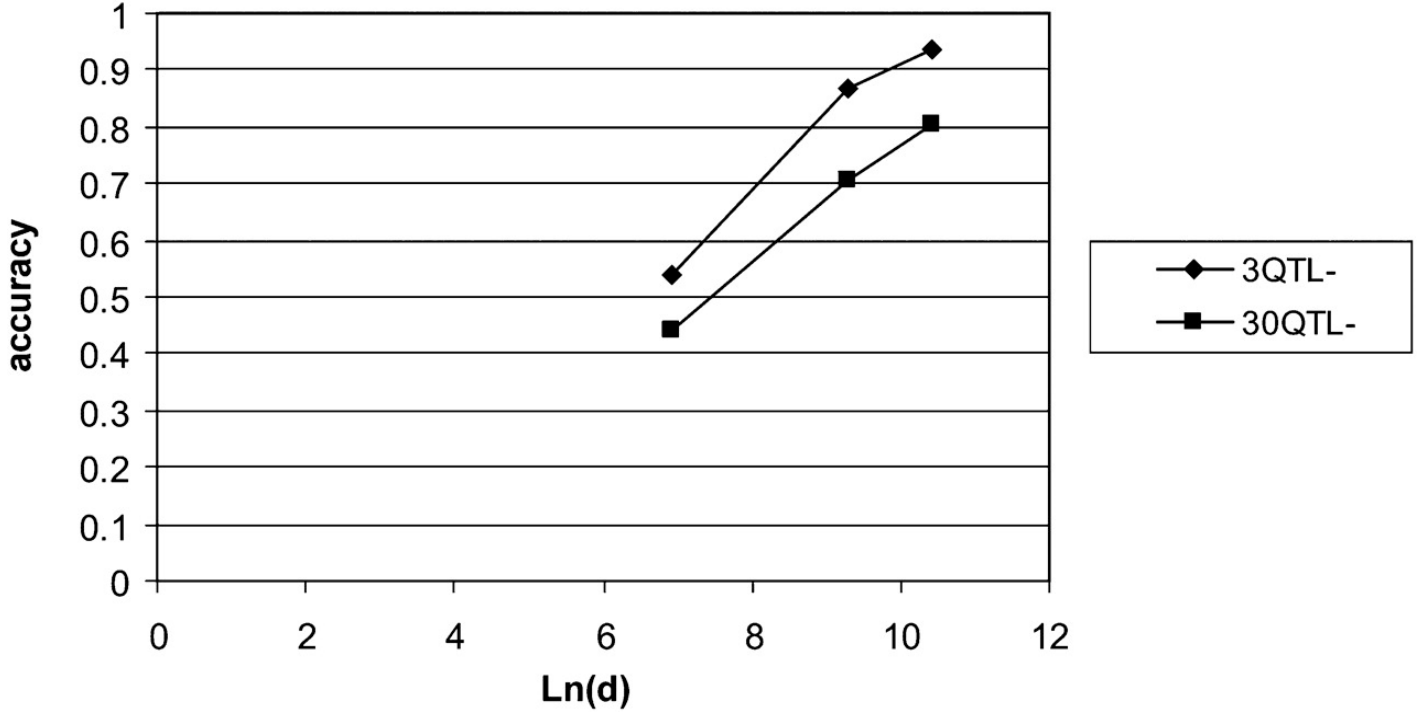
EAAP 2016

Belfast

Overview

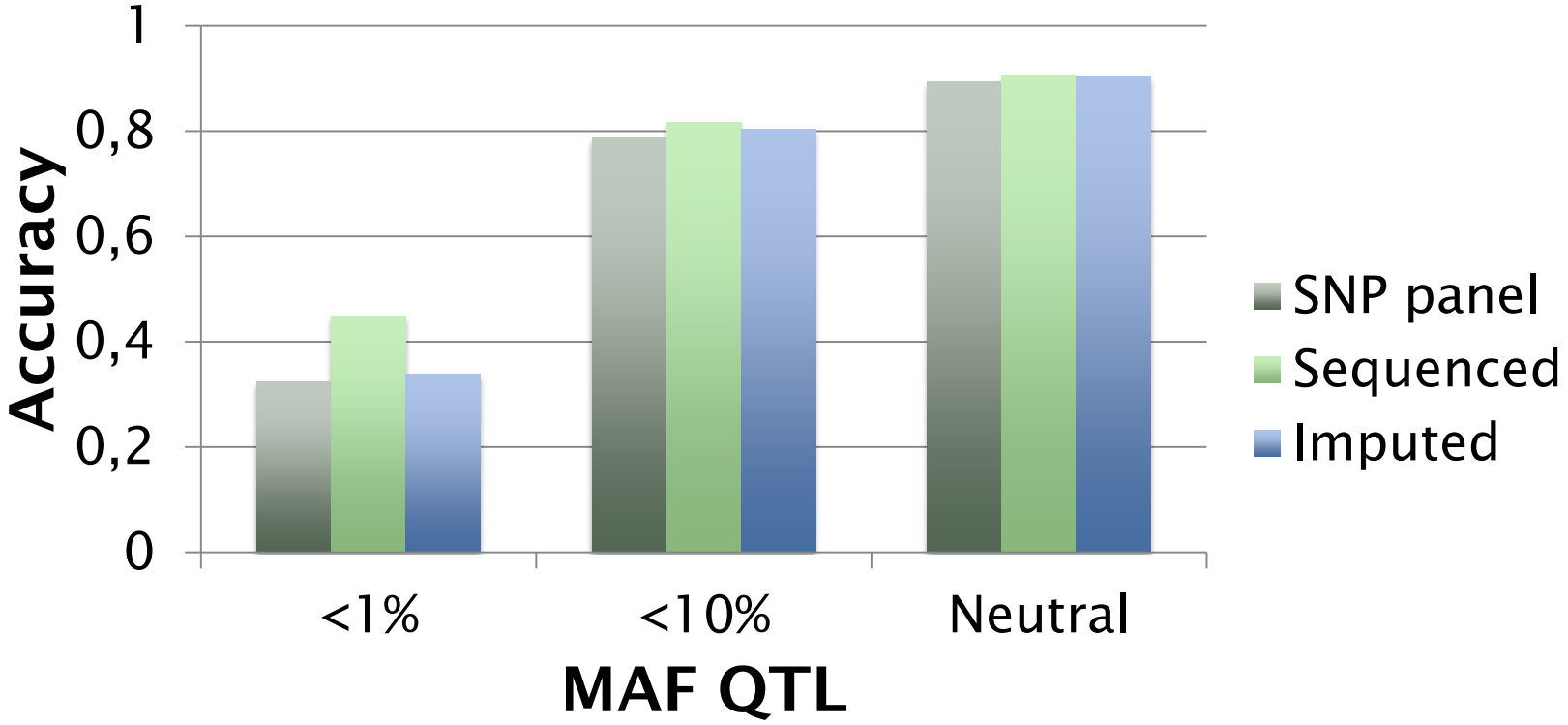
- Introduction: Simulation studies
- Material & Methods
- Results:
  - 50k GenSel
  - 50k and Sequence gbcpp
  - LD-pruned sequence gbcpp
  - Other densities: missense variants and top variants from GWAS
- Conclusion

Genomic prediction: Simulation studies



Meuwissen and Goddard (2010) *Genetics*

Another simulation study



Modified after Druet et al. 2014 *Heredity*

And now with real data ...

We have...

- ... Genotypes of 23,000 Brown Swiss (imputed to sequence level)
- ... 3 traits investigated

	Non-return rate 56 heifer (NRH)	Somatic cell score (SCS)	Stature (STA)
Reference	2,018	4,786	5,294
Validation	240	560	596

Algorithms

- Softwares/ Algorithms
  - BayesC implemented in GenSel (R. Fernando and D. Garrick, 2008)

$$y = 1'\mu + X\beta + e$$

$$\beta_i \begin{cases} = 0 \text{ with probability } \pi \\ \sim N(0, \sigma^2_{SNP}) \text{ with probability } (1-\pi) \end{cases}$$

Algorithms

- Softwares/ Algorithm
  - BayesC-like implemented in gbcpp (T. Meuwissen)

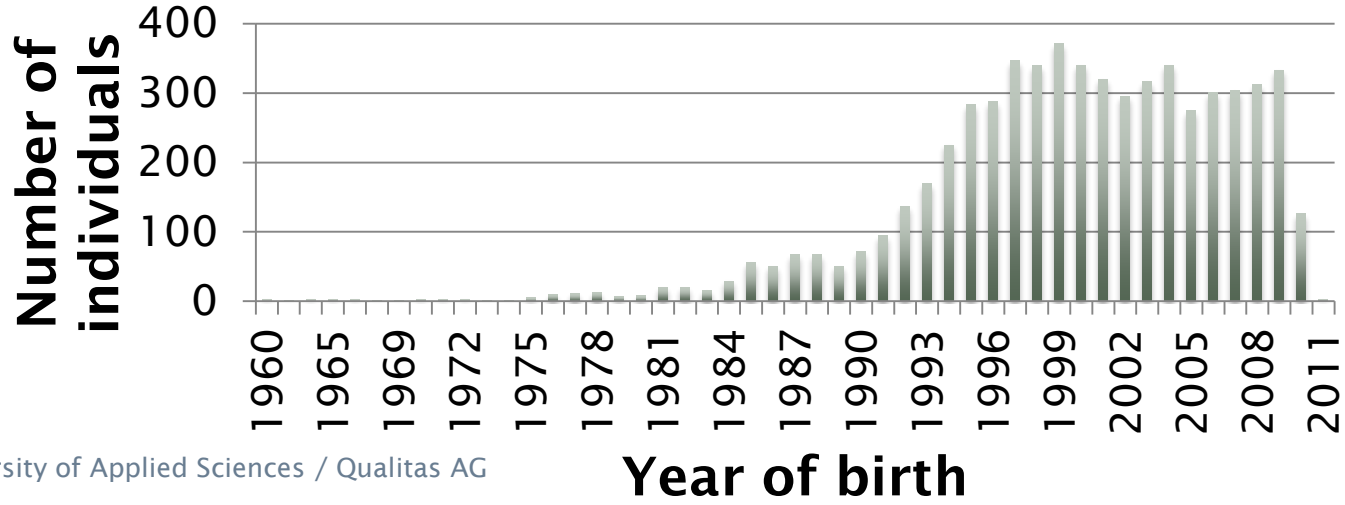
$$y = 1'\mu + Zg + X\beta + e$$

$$g \sim N(0, G\sigma_g^2)$$

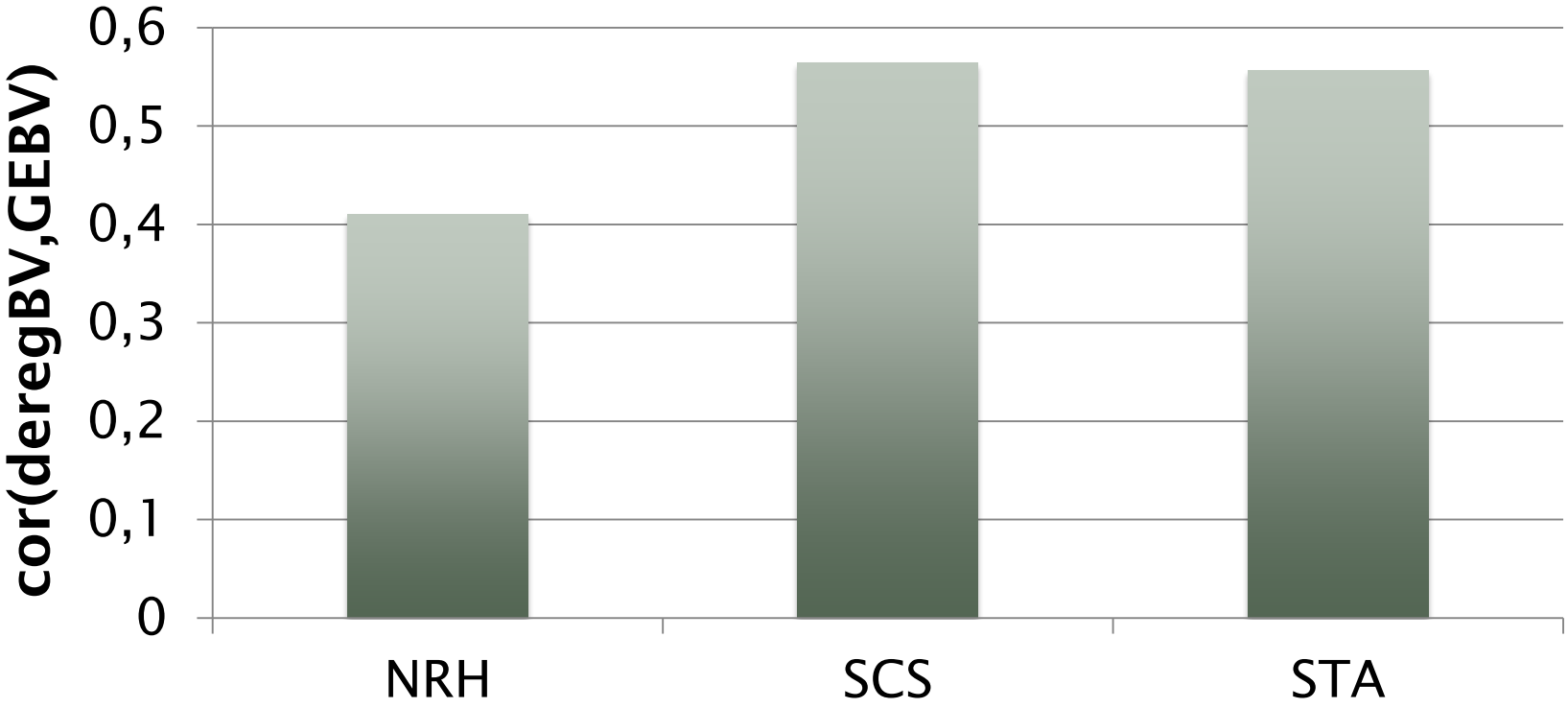
$$\beta_i \begin{cases} = 0 \text{ with probability } \pi \\ \sim N(0, \sigma_{SNP}^2) \text{ with probability } (1-\pi) \end{cases}$$

Phenotypes

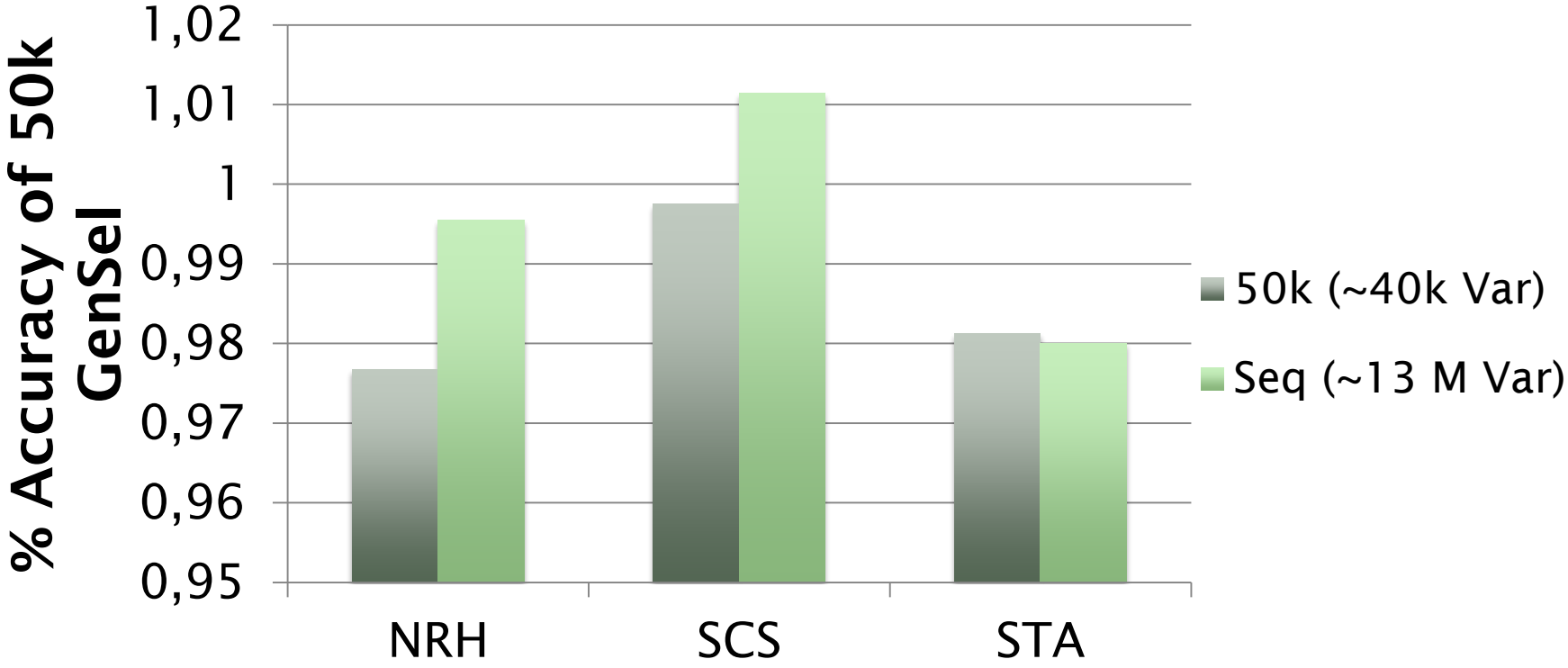
- Minimal reliability
- Deregressed BV (Garrick et al., 2009)
- Youngest bulls as validation
- Prediction accuracy: cor(deregressed BV,GEBV)



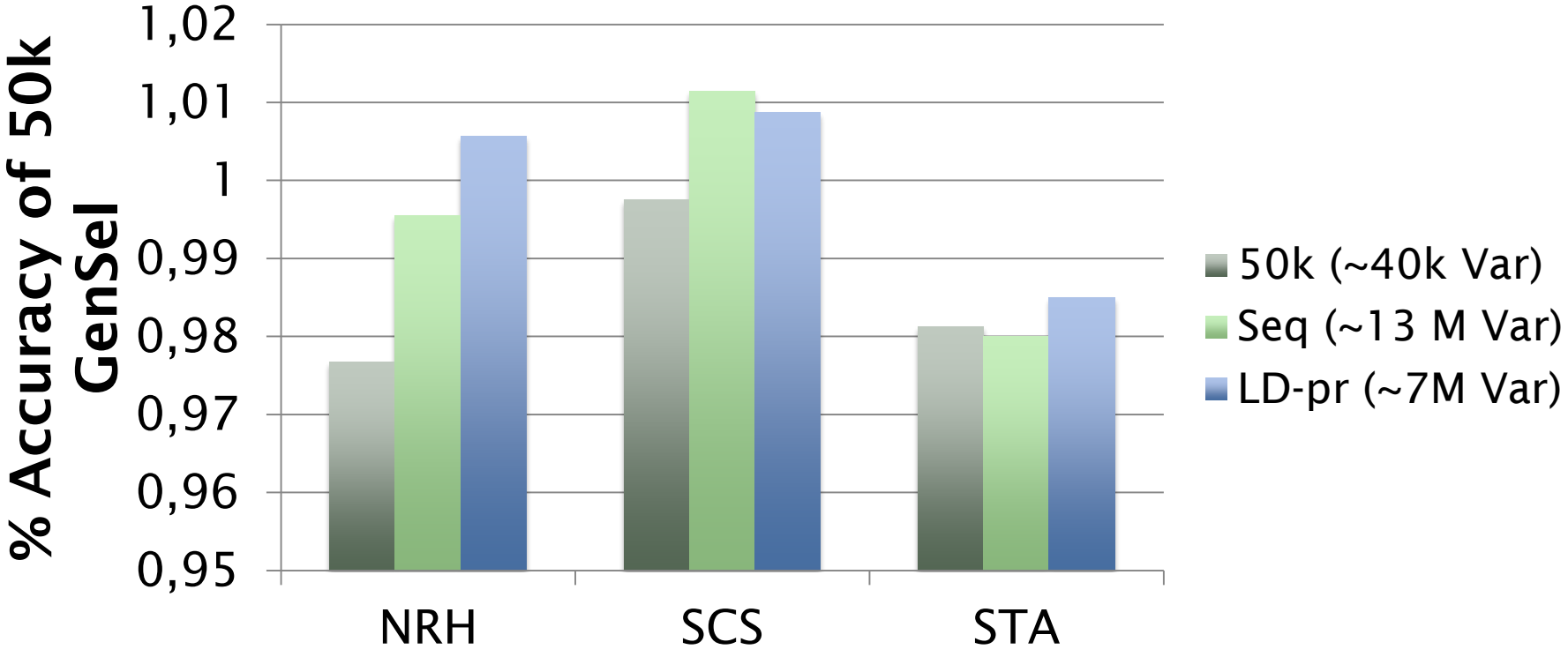
Results -50k GenSel



Results –Sequence data with gbcpp



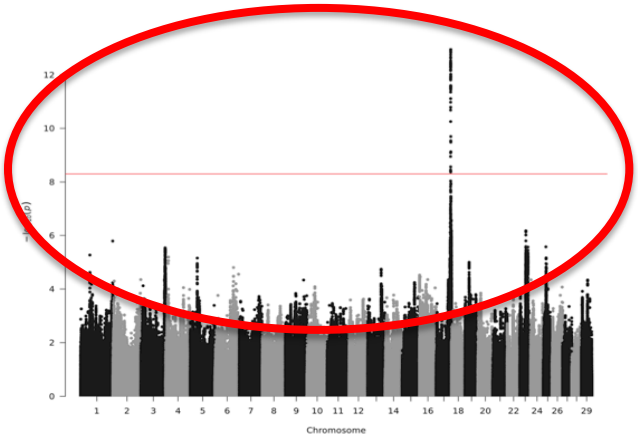
Results – LD-pruned sequence data with gbcpp



Results –Other densities

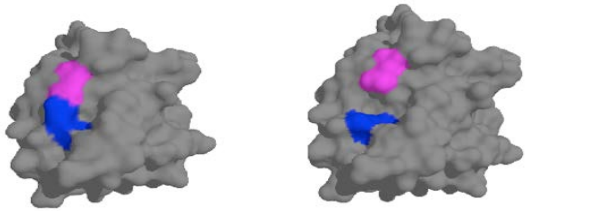
„Top“

- GWAS results
- 50,000 Variants



„Missense“

- Missense variants in protein coding regions
- About 35,000 Variants



https://en.wikipedia.org/wiki/LMNA#/media/File:LMNA_protein_%281ifr%29_mutation_R527L_PMID_22549407_surface_and_cartoon.png

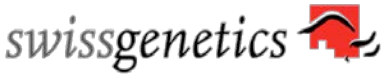
Results - Other densities

Density	NRH	SCS	STA
Top GBCPP	-0.15	-0.08	-0.10
Miss GBCPP	-0.06	-0.05	-0.02
Top GenSel	-0.11	-0.07	-0.09
Miss GenSel	0.01	-0.01	0.01

Conclusion

- No improvement using sequencing data
- Also other densities do not improve the accuracy
- Currently sequence data is not an option for routine evaluation

Acknowledgements



Thank you for your attention