Genomic inbreeding estimates with imputed genotypes in Italian Holstein dairy cattle

Christos Dadousis[1], Michela Ablondi[1], Jan-Thijs van Kaam[2], Alberto Sabbioni[1]

[1]*Department of Veterinary Science, University of Parma, Via del Taglio 10, 43126 Parma, Italy*

[2]*ANAFIBJ - Associazione Nazionale Allevatori della Razza Frisona, Bruna e Jersey Italiana, Via Bergamo 292, 26100 Cremona, Italy*

Introduction

BMC Genomics

METHODOLOGY ARTICLE   Open Access

How to study runs of homozygosity using PLINK? A guide for analyzing medium density SNP data in livestock and pet species

R. Meyermans[†], W. Gorssen[†], N. Buys and S. Janssens[*]

RESEARCH ARTICLE

The value of genomic relationship matrices to estimate levels of inbreeding

Beatriz Villanueva[1*], Almudena Fernández[1], María Saura[1], Armando Caballero[2], Jesús Fernández[1], Elisabeth Morales-González[1], Miguel A. Toro[3] and Ricardo Pong-Wong[4]

frontiers in Genetics

animal

Comparative evaluation of genomic inbreeding parameters in seven commercial and autochthonous pig breeds

G. Schiavo[1], S. Bovo[1], F. Bertolini[2], S. Tinarelli[1,3], S. Dall'Olio[1], L. Nanni Costa[1], M. Gallo[3] and L. Fontanesi[1†]

[1]Division of Animal Sciences, Department of Agricultural and Food Sciences, University of Bologna, Viale G. Fanin 46, 40127 Bologna, Italy; [2]National Institute of Aquatic Resources, Technical University of Denmark, Kemitorvet, Building 202, 2800 Kongens Lyngby, Denmark; [3]Associazione Nazionale Allevatori Suini, Via Nizza 53, 00198 Roma, Italy

frontiers in Genetics

Controlling Coancestry and Thereby Future Inbreeding by Optimum-Contribution Selection Using Alternative Genomic-Relationship Matrices

G. T. Gebregiwergis[1*], Anders C. Sørensen[2], Mark Henryon[3,4] and Theo Meuwissen[1]

[1]Department of Animal and Aquaculture Sciences, Norwegian University of Life Sciences, Ås, Norway, [2]Department of Molecular Biology and Genetics, Aarhus University, Aarhus, Denmark, [3]Seges, Copenhagen, Denmark, [4]School of Agriculture and Environment, University of Western Australia, Crawley, WA, Australia

GSE Genetics Selection Evolution

RESEARCH   Open Access

Detecting inbreeding depression for reproductive traits in Iberian pigs using genome-wide data

María Saura[1*], Almudena Fernández[1], Luis Varona[2], Ana I Fernández[1], Maria Ángeles R de Cara[3], Carmen Barragán[1] and Beatriz Villanueva[1]

Management of Genetic Diversity in the Era of Genomics

Theo H. E. Meuwissen[1*], Anna K. Sonesson[2], Gebreyohans Gebregiwergis[1] and John A. Woolliams[3]

[1]Department of Animal and Aquacultural Sciences, Norwegian University of Life Sciences, Ås, Norway, [2]The Rosin Institute and R(D)SVS, The University of Edinburgh, Edinburgh, United Kingdom

Human Heredity

Population Aspects of Consanguinity

Inbreeding Coefficient Estimation with Dense SNP Data: Comparison of Strategies and Application to HapMap III

Steven Gazal[a, c]   Mourad Sahbatou[d]   Hervé Perdry[b, e]   Sébastien Letort[f, g]   Emmanuelle Génin[f, g]   Anne-Louise Leutenegger[a, h]

[a]Genetic Variability and Human Diseases, Inserm, U946, [b]UMR 669 at [c]Université Paris Sud, and [d]Fondation Jean Dausset CEPH, Paris, [e]Inserm, U669, Villejuif, [f]Génétique, Génomique Fonctionnelle et Biotechnologies, Inserm, U1078, and [g]Centre Hospitalier Régional et Universitaire de Brest, Brest, and [h]Institut Universitaire d'Hématologie, UMR 946, Université Paris-Diderot, Paris, France

Detection and quantification of inbreeding depression for complex traits from SNP data

Loic Yengo[a,1], Zhihong Zhu[a], Naomi R. Wray[a,b], Bruce S. Weir[c], Jian Yang[a,b], Matthew R. Robinson[a,d], and Peter M. Visscher[a,b,1]

[a]Institute for Molecular Bioscience, The University of Queensland, Brisbane, QLD 4072, Australia; [b]Queensland Brain Institute, The University of Queensland, Brisbane, QLD 4072, Australia; [c]Department of Biostatistics, University of Washington, Seattle, WA 98195; and [d]Department of Computational Biology, University of Lausanne, Lausanne, CH-1015, Switzerland

Edited by Andrew G. Clark, Cornell University, Ithaca, NY, and approved June 21, 2017 (received for review December 22, 2016)

PNAS

Genomic inbreeding debate

Introduction

❑ Inbreeding estimates derived from **pedigree** data (F_{ped})
  - Question on pedigree quality
  - IBD theory

❑ **Genomic** data (mainly in the form of SNP) is expected to be more precise on the estimation of the inbreeding level. However, not straightforward…:
  - IBS theory
  - Different algorithms [4 plink, ≥ 3 GRM, ROH (various definitions of ROH)]
  - Different SNP quality control (call rate, MAF, HWE, LD)
  - Different genotypic SNP panels (SNP density, commercial vs. private)
  - SNP imputation (different algorithms and strategies) → the rule nowadays
  - …

Objectives

❑ Compare different algorithms for estimating genomic inbreeding coefficients with imputed SNP data

❑ Compare genotyped vs. imputed SNP genomic inbreeding coefficients

Material and Methods

ANAFIBJ

95,540 cows!

84,445 imputed SNP

30 SNP panels!

SNP/panel
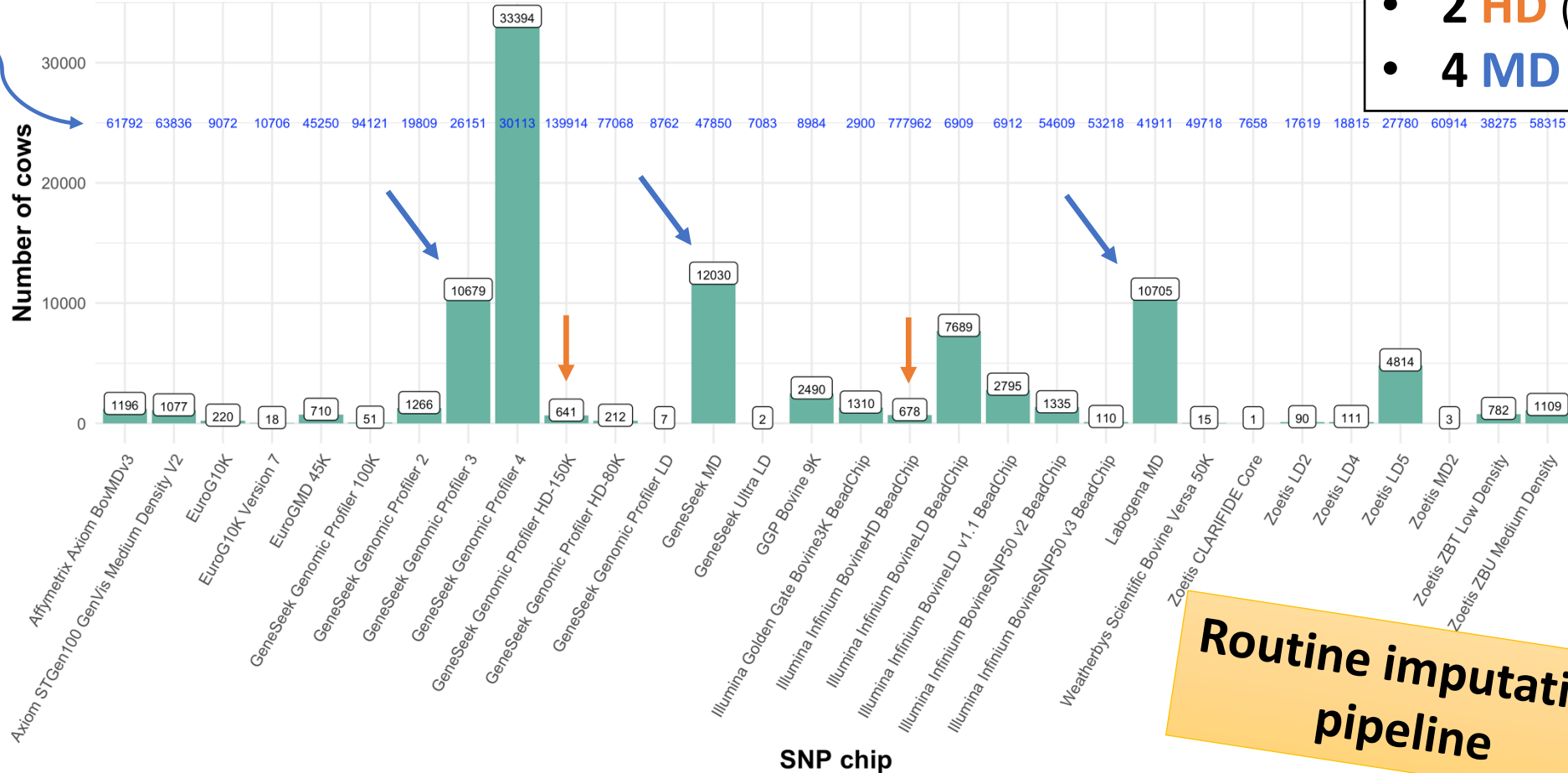
Select:
- **2 HD** (~600 cows)
- **4 MD** (~10-30k cows)

Routine imputation pipeline

Number of cows vs SNP chip

Values above bars (blue): 61792 63836 9072 10706 45250 94121 19809 26151 30113 139914 77068 8762 47850 7083 8984 2900 777962 6909 6912 54609 53218 41911 49718 7658 17619 18815 27780 60914 38275 58315

Bar labels: 1196 1077 220 18 710 51 1266 10679 33394 641 212 7 12030 2 2490 1310 678 7689 2795 1335 110 10705 15 1 90 111 4814 3 782 1109

SNP chip categories (x-axis):
Affymetrix Axiom BovMDv3, Axiom STGen100 GenVis Medium Density V2, EuroG10K, EuroG10K Version 7, EuroGMD 45K, GeneSeek Genomic Profiler 100K, GeneSeek Genomic Profiler 2, GeneSeek Genomic Profiler 3, GeneSeek Genomic Profiler 4, GeneSeek Genomic Profiler HD-150K, GeneSeek Genomic Profiler HD-80K, GeneSeek Genomic Profiler LD, GeneSeek MD, GeneSeek Ultra LD, GGP Bovine 9K, Illumina Golden Gate Bovine3K BeadChip, Illumina Infinium BovineHD BeadChip, Illumina Infinium BovineLD BeadChip, Illumina Infinium BovineLD v1.1 BeadChip, Illumina Infinium BovineSNP50 v2 BeadChip, Illumina Infinium BovineSNP50 v3 BeadChip, Labogena MD, Weatherbys Scientific Bovine Versa 50K, Zoetis CLARIFIDE Core, Zoetis LD2, Zoetis LD4, Zoetis LD5, Zoetis MD2, Zoetis ZBT Low Density, Zoetis ZBU Medium Density

Material and Methods

❑ Methodologies:

- Fped (*optiSel*, *R*)

  > Allelic frequency dependent

- GRM
  - VanRaden (method 1 & 3; Fgrm & Fgrm2, respectively)

  > Allelic frequency independent
  > Regressed on Fped

- F *plink* (plink imports the algorithms from GCTA)
  - F method of moments based on observed and expected autosomal homozygous genotype counts for each sample [(obs.hom.count-exp.count)/(total obs-exp.count)])

    > (--het flag)

  - F_{hat1} (usual variance-standardized relationship minus 1, VanRadens' 2nd)
  - F_{hat2} ($\cong$ F)
  - F_{hat3} (based on the correlation between uniting gametes)

    > (--ibc flag)

- FROH (consecutiveRUNS.run, *detectRUNS*, R)
  - ROH definition: minSNP = 20, maxGap = 10^6, minLengthBps = 10^6, maxOppRun = 1, maxMissRun = 1

Confusion in the literature…

Villanueva et al., 2021
PMID: 33933002

Table 1 Summary of the names given to different genomic inbreeding coefficients in the literature

Results and Discussion

Illumina Infinium BovineHD BeadChip (n = 678, nSNP=79,900 – HD)

genotyped

imputed

- Consistent

Results and Discussion

Illumina Infinium BovineHD BeadChip (n = 678, nSNP=79,900 – HD)
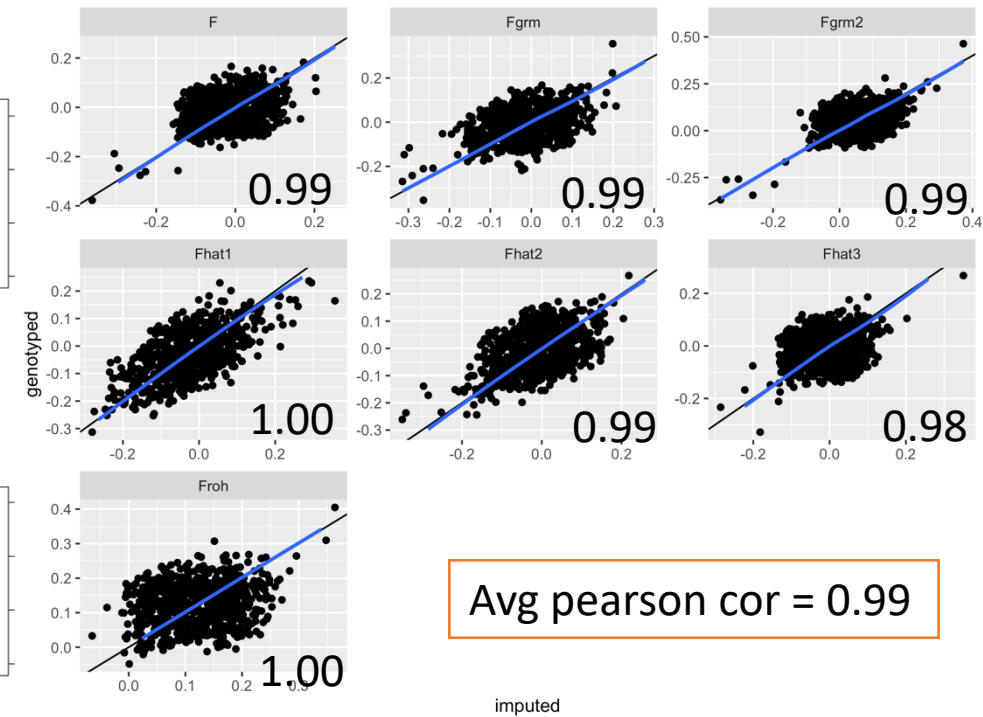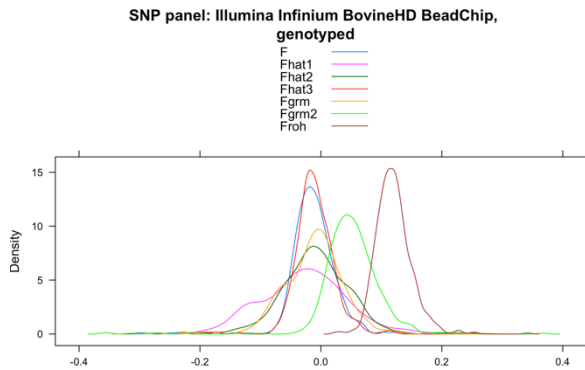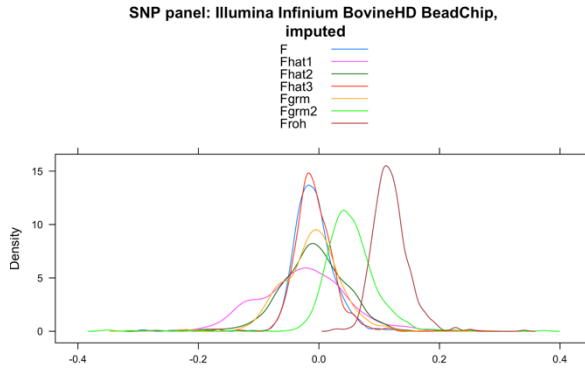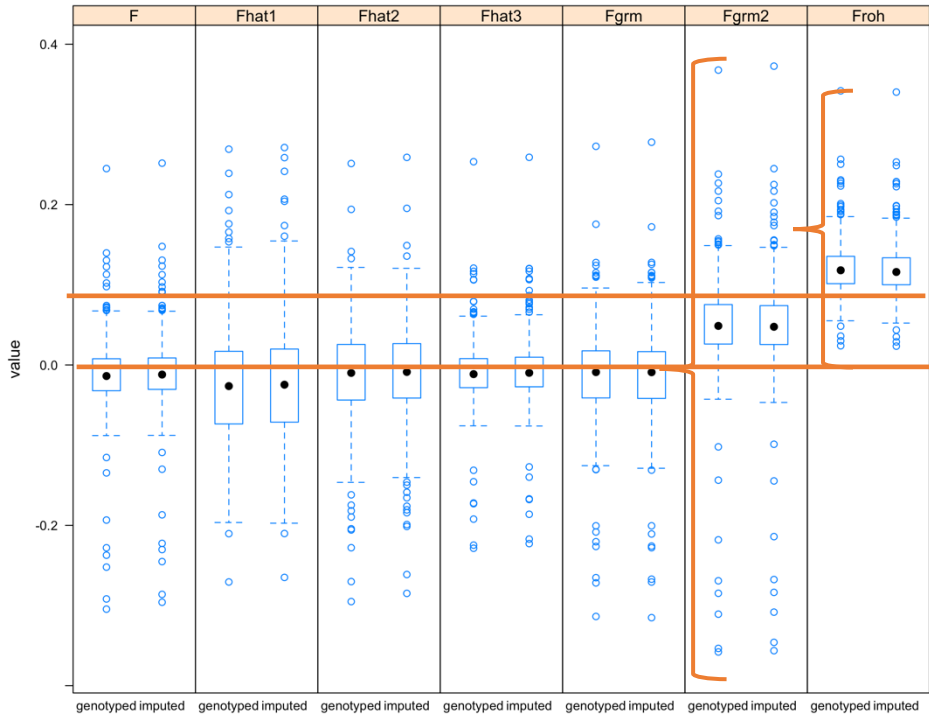
- Froh highest mean
- Froh>0
- All but Froh [-,+]
- Fgrm2 highest variability
- Froh lowest variability
- Comparable variability between genotyped-imputed/method

Results and Discussion

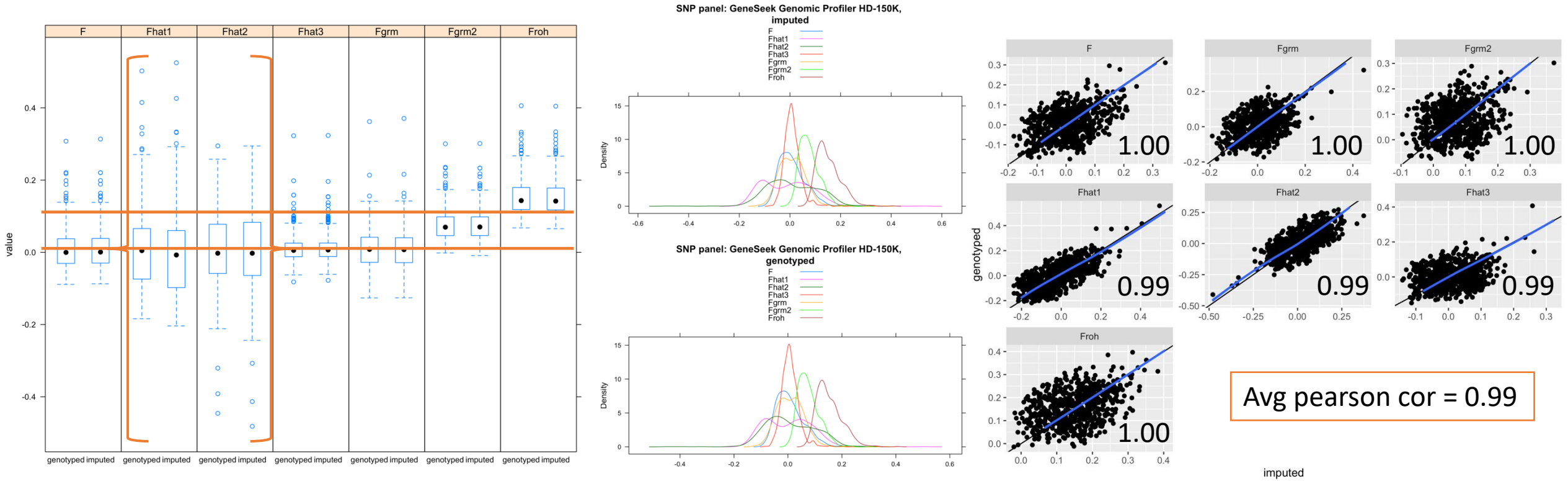## GeneSeek Genomic Profiler HD-150K (n=641, nSNP=77,085 – HD)

genotyped
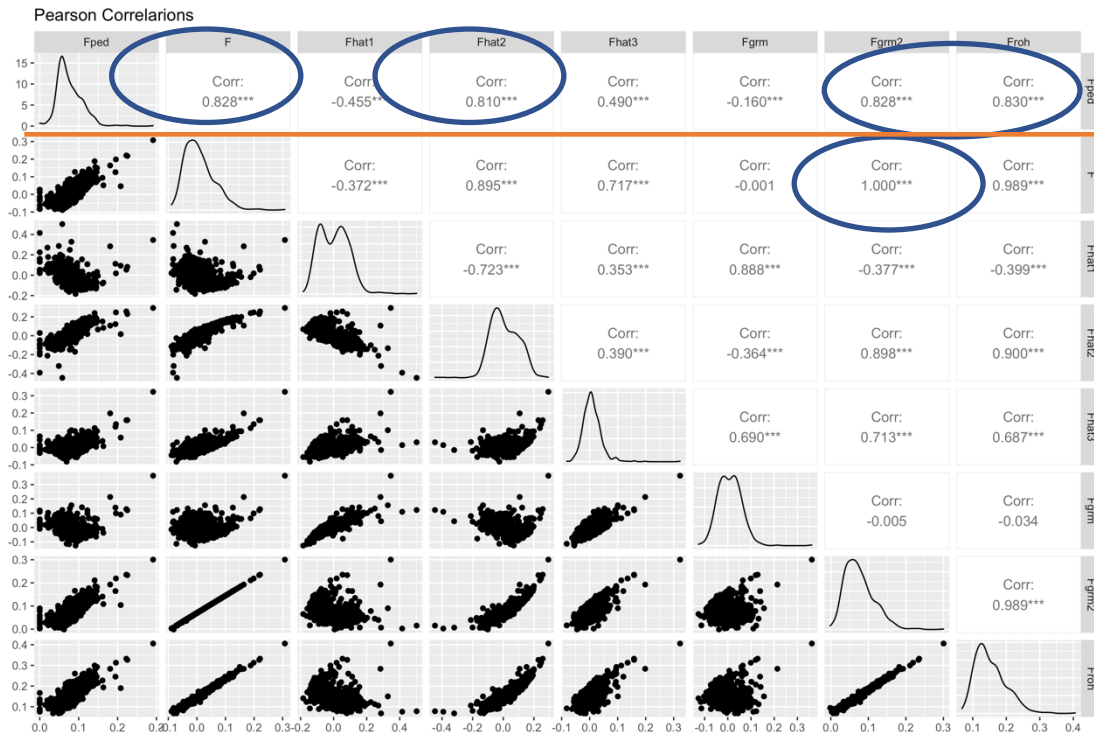
imputed
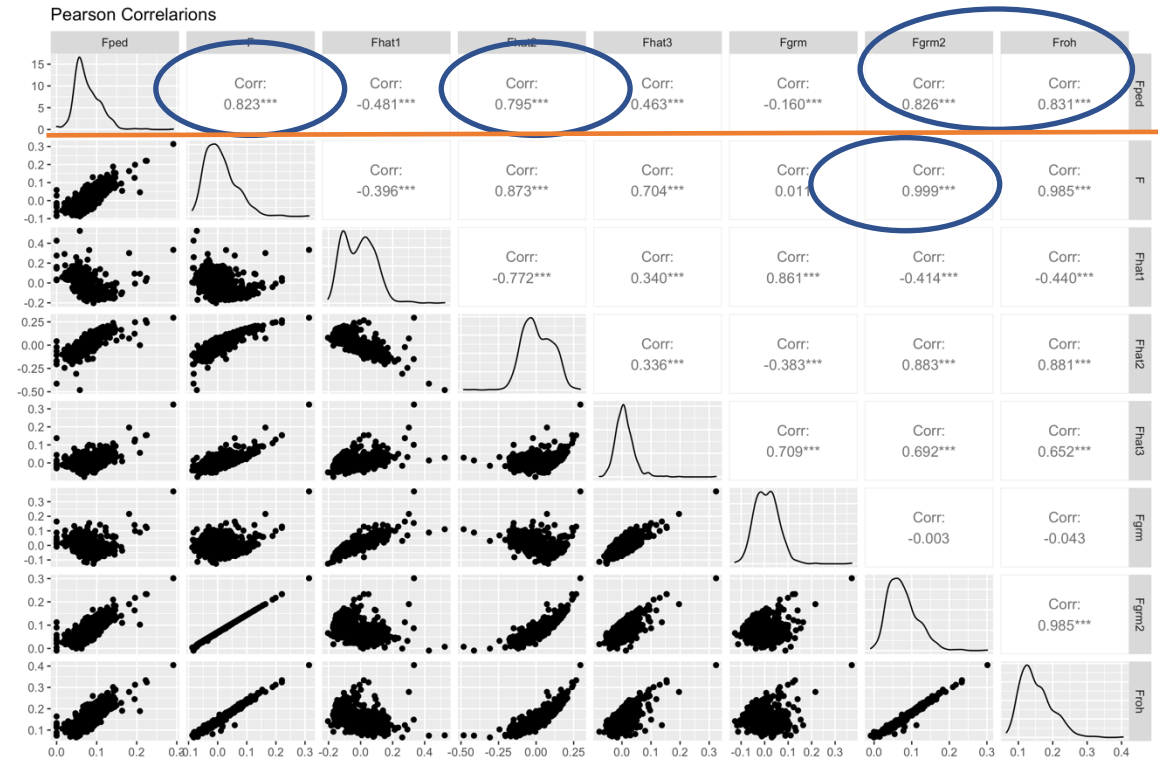
- Consistent
- F ≅ Fgrm2
- Corr to Fped higher compared to illumina

Results and Discussion

GeneSeek Genomic Profiler HD-150K (n=641, nSNP=77,085 – HD)

- Comparable variability between genotyped-imputed/method
- Variability across methods

Results and Discussion

GeneSeek Profiler 3 (n=10,679, nSNP=13,870 – MD)

genotyped

imputed
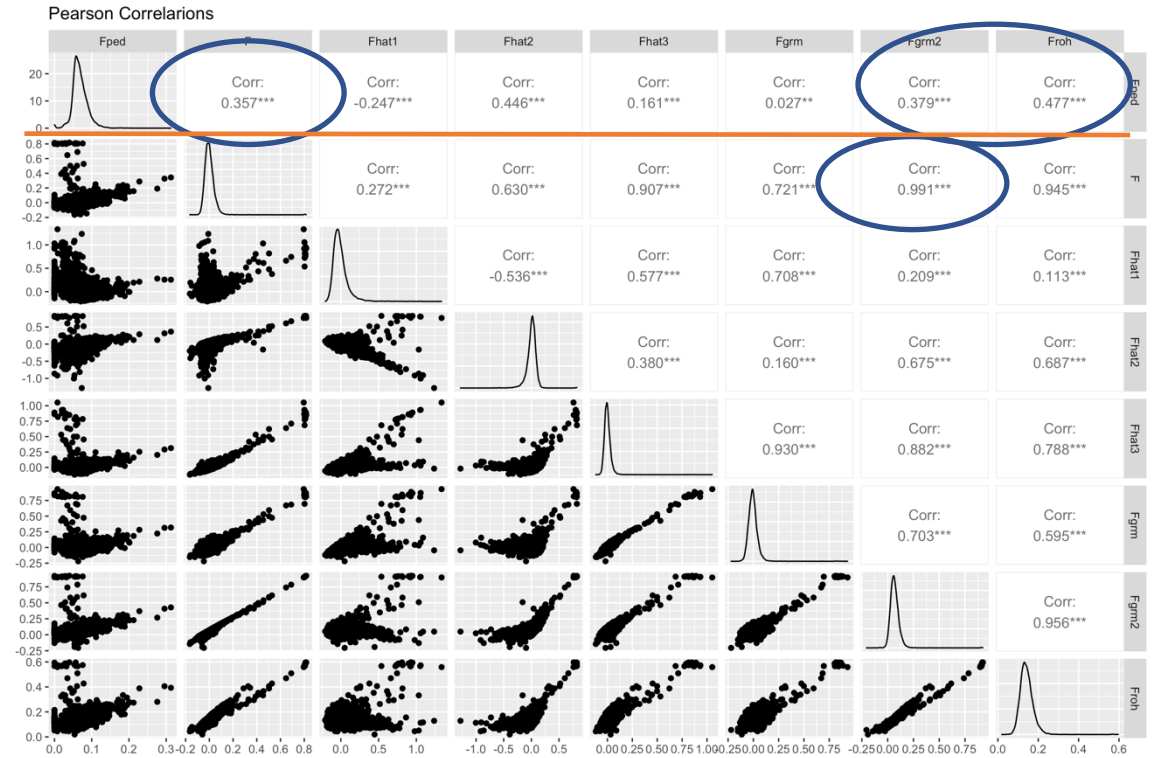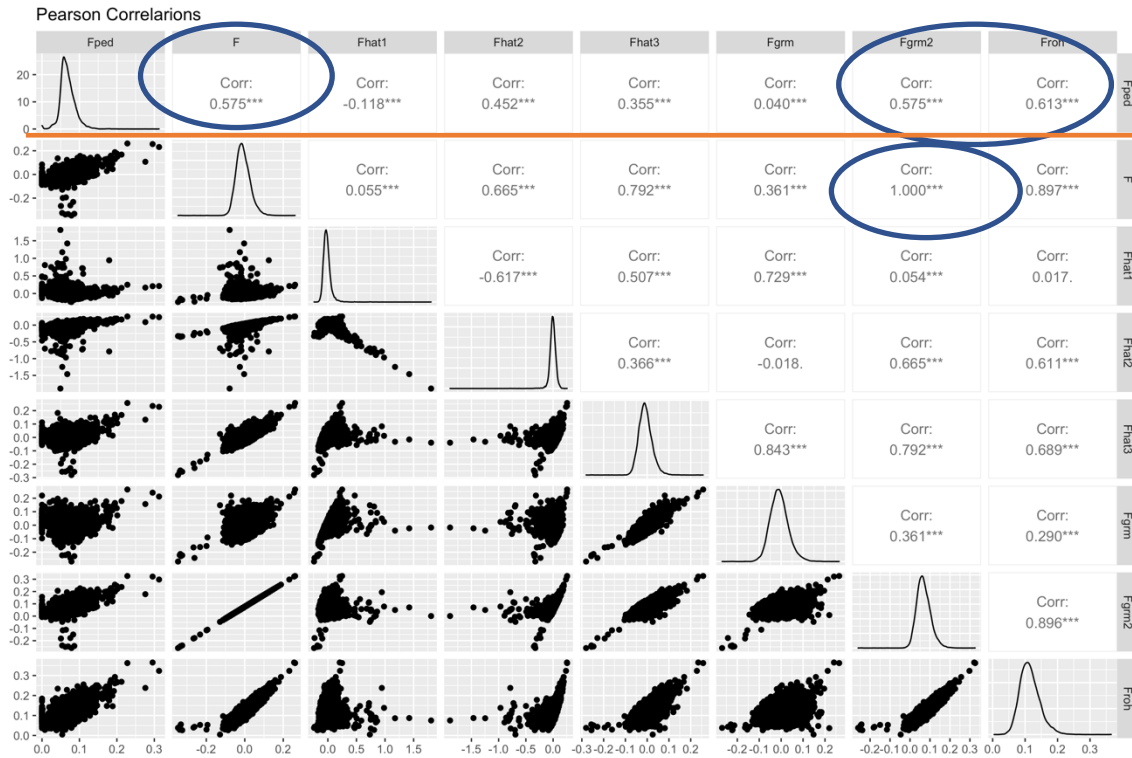
- Corr to Fped higher in genotyped

Results and Discussion

GeneSeek Profiler 3 (n=10,679, nSNP=13,870 – MD)

Avg pearson cor = 0.65
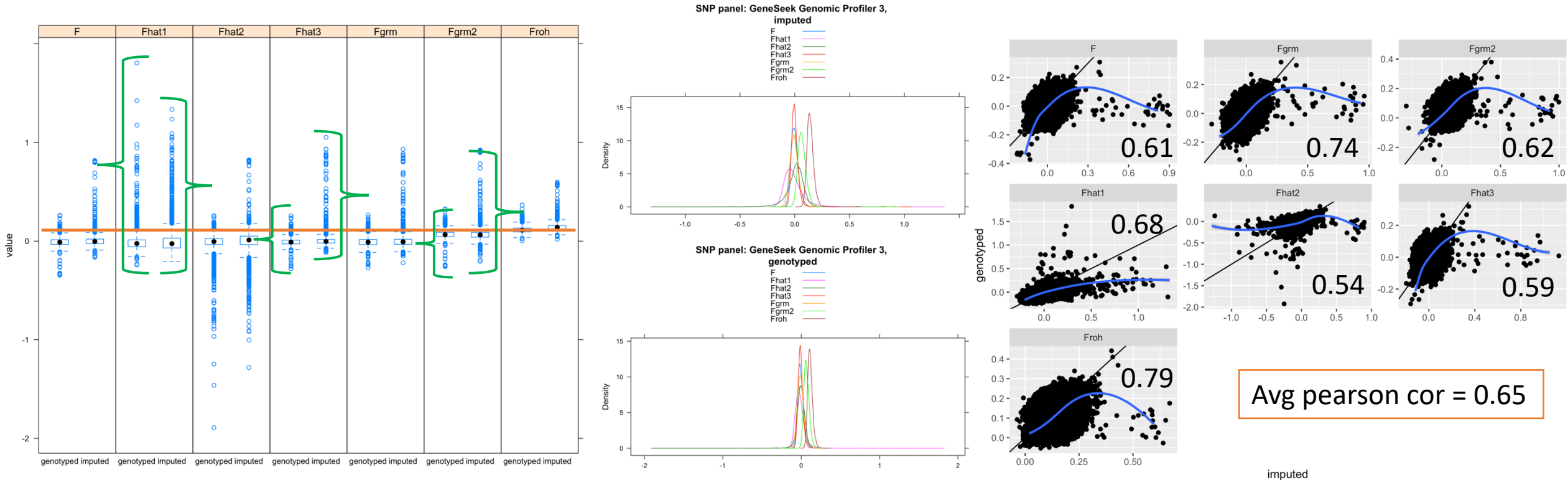
- Variability between genotyped-imputed
- For all but Fhat1 the imputed had higher variability than the genotyped

Results and Discussion

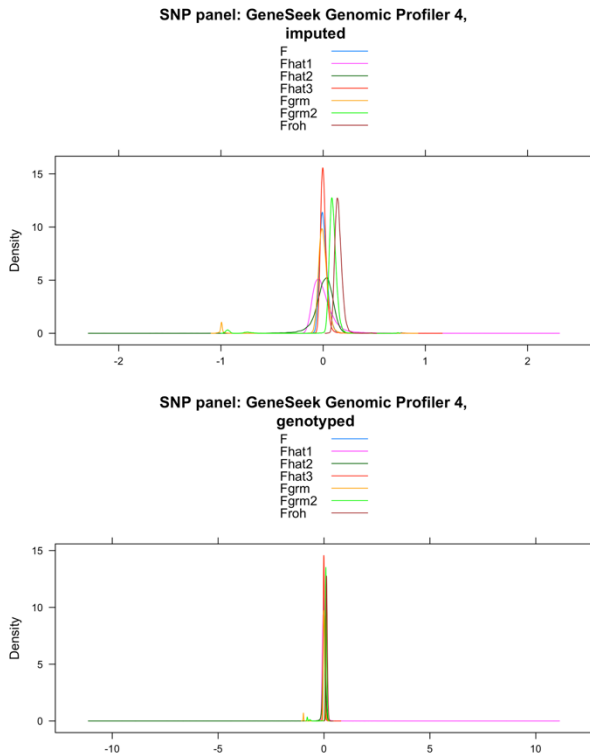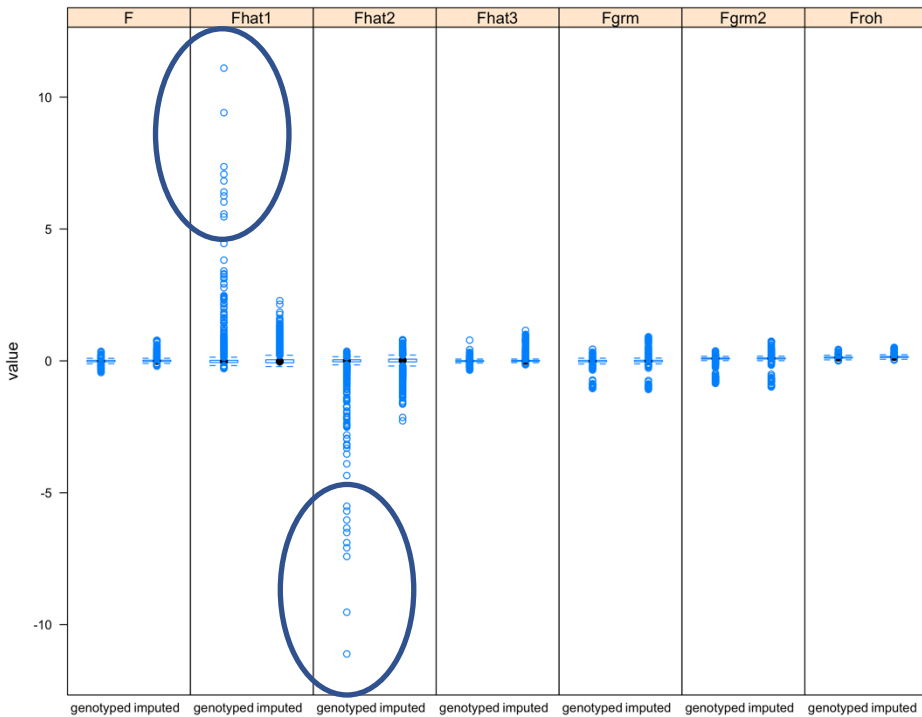**GeneSeek Profiler 4 (n=33,394, nSNP=16,862 – MD)**

genotyped

imputed
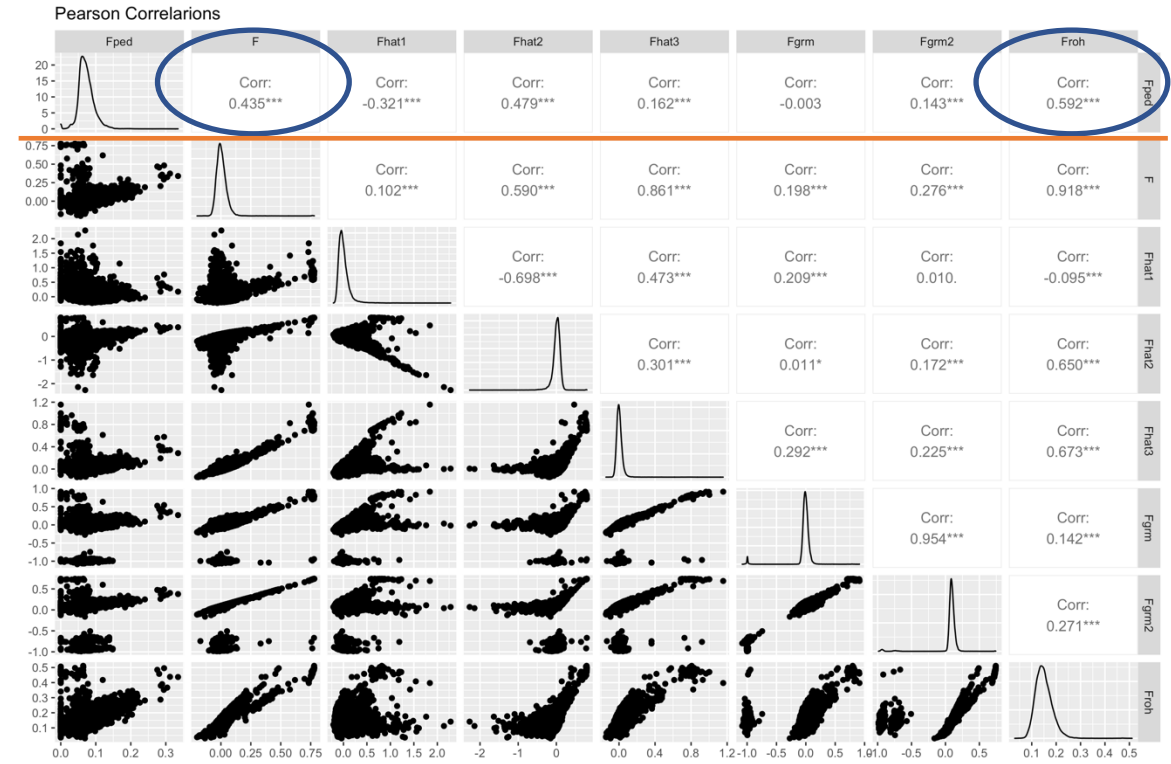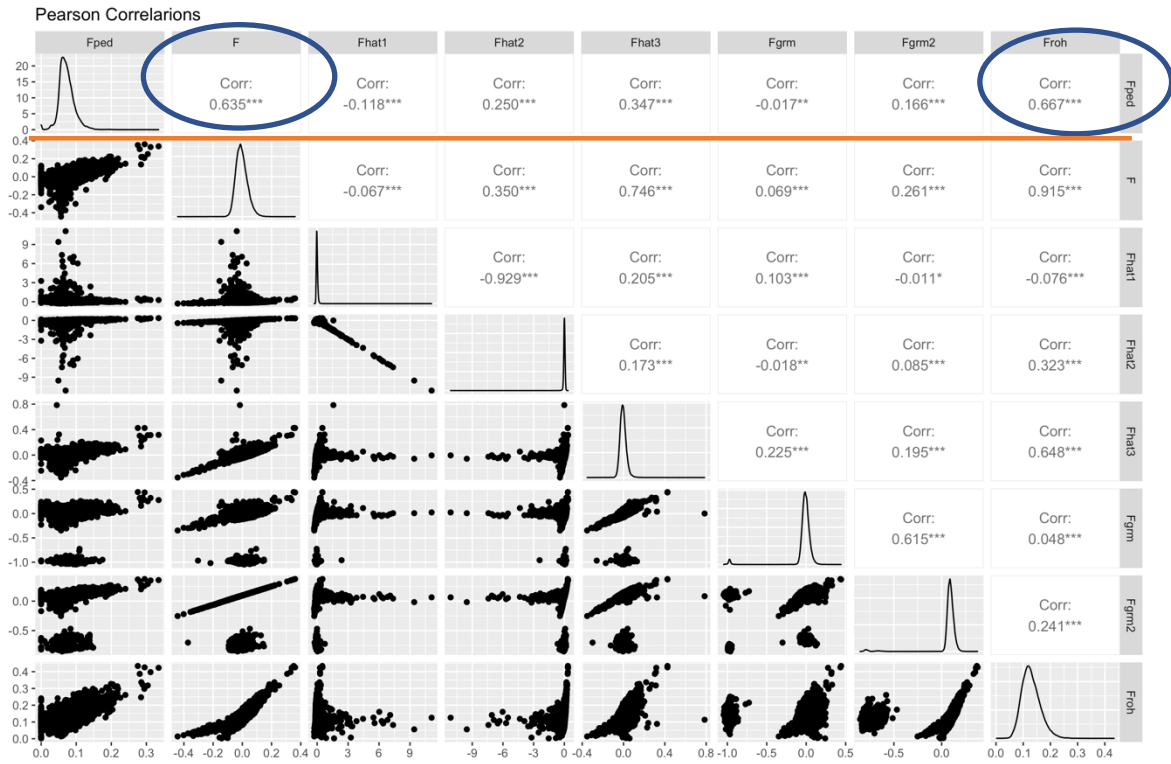
- Corr to Fped higher in genotyped … intermediate results

Results and Discussion

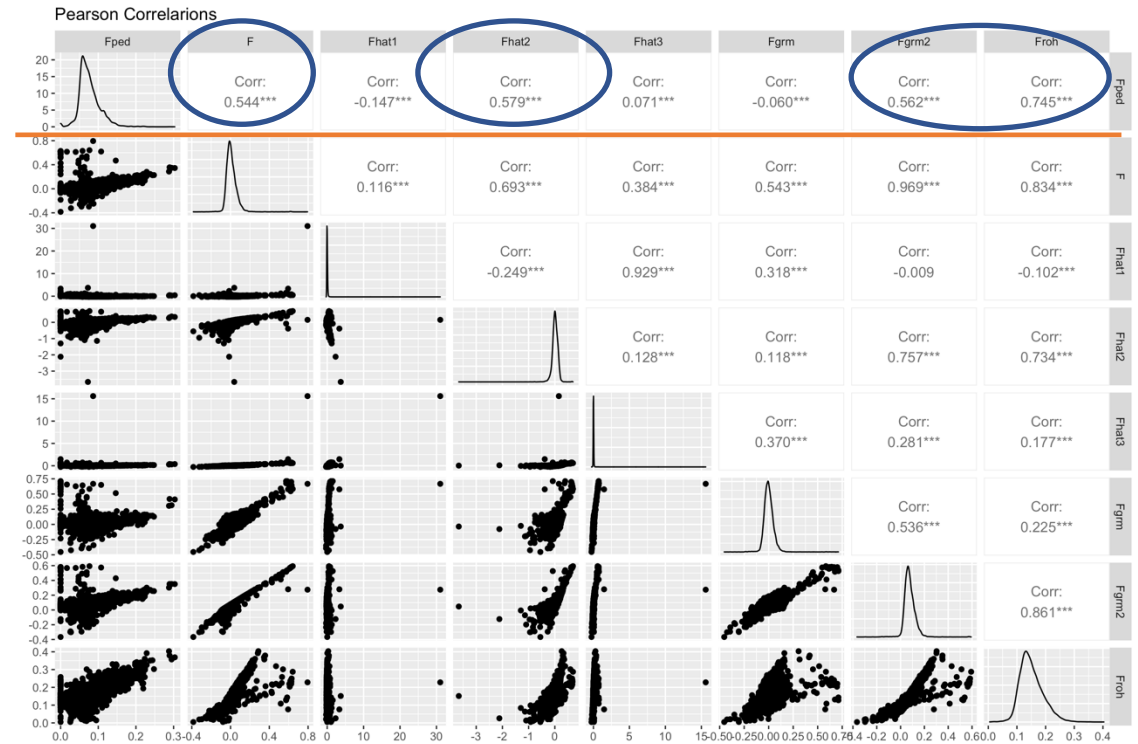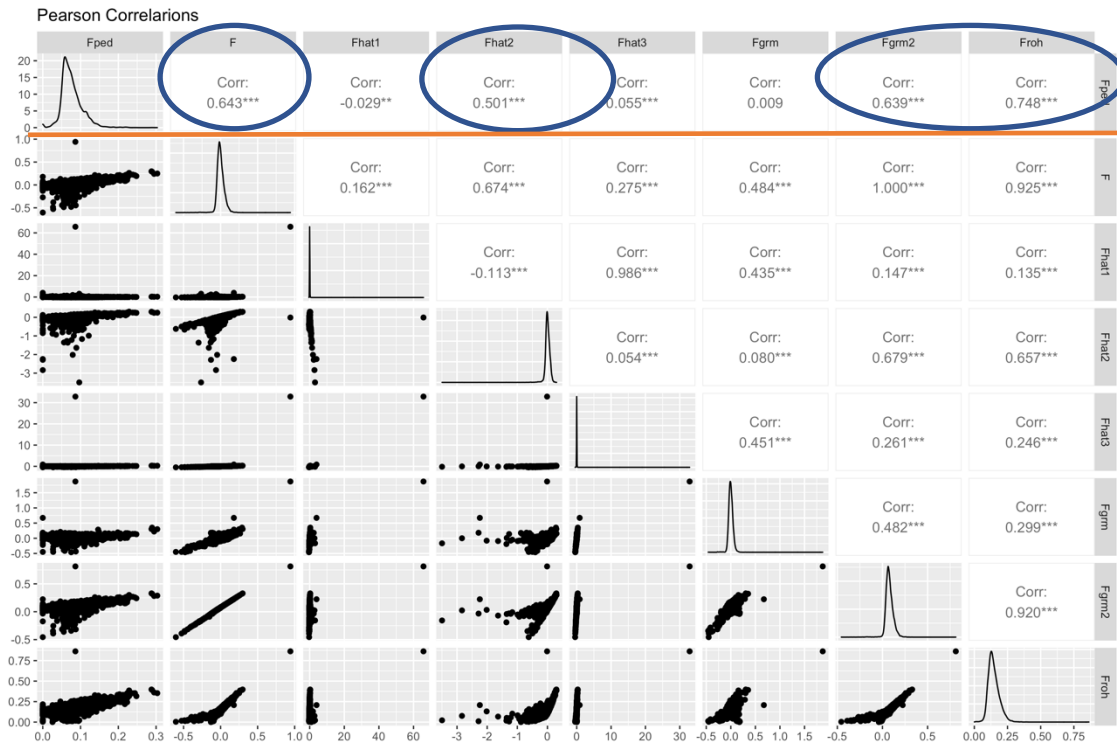**GeneSeek Profiler 4 (n=33,394, nSNP=16,862 – MD)**
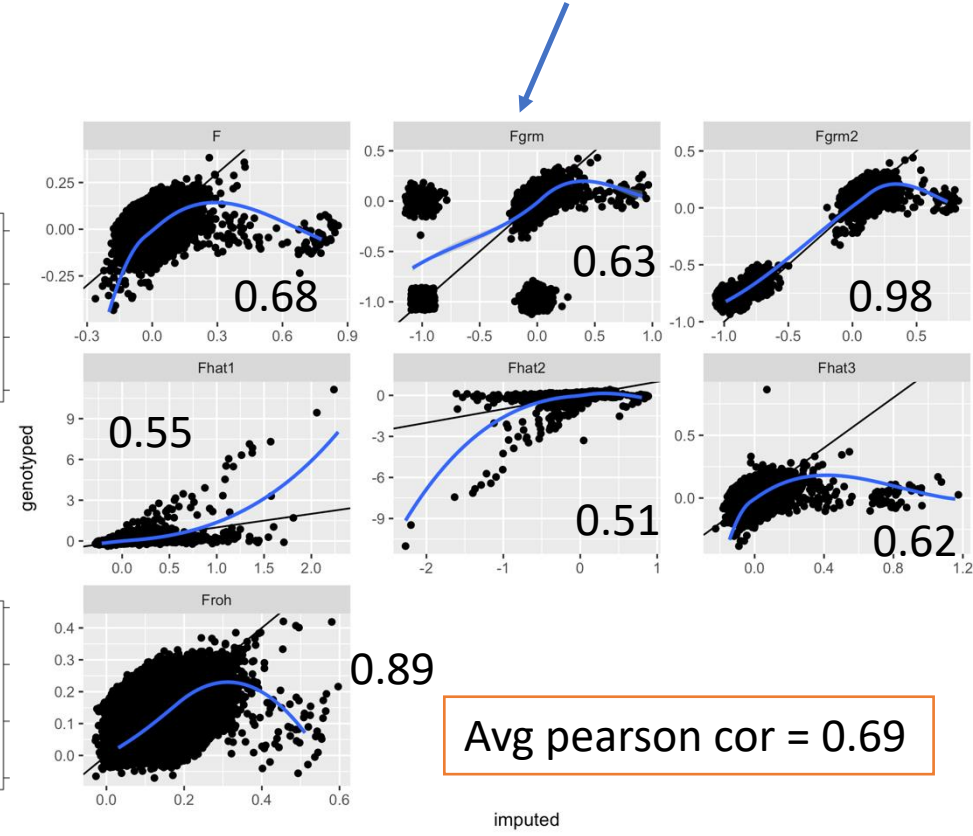
- For all but Fhat1 and Fhat2 the imputed had higher variability than the genotyped
- Extreme values with genotyped
- Subgrouping

Results and Discussion
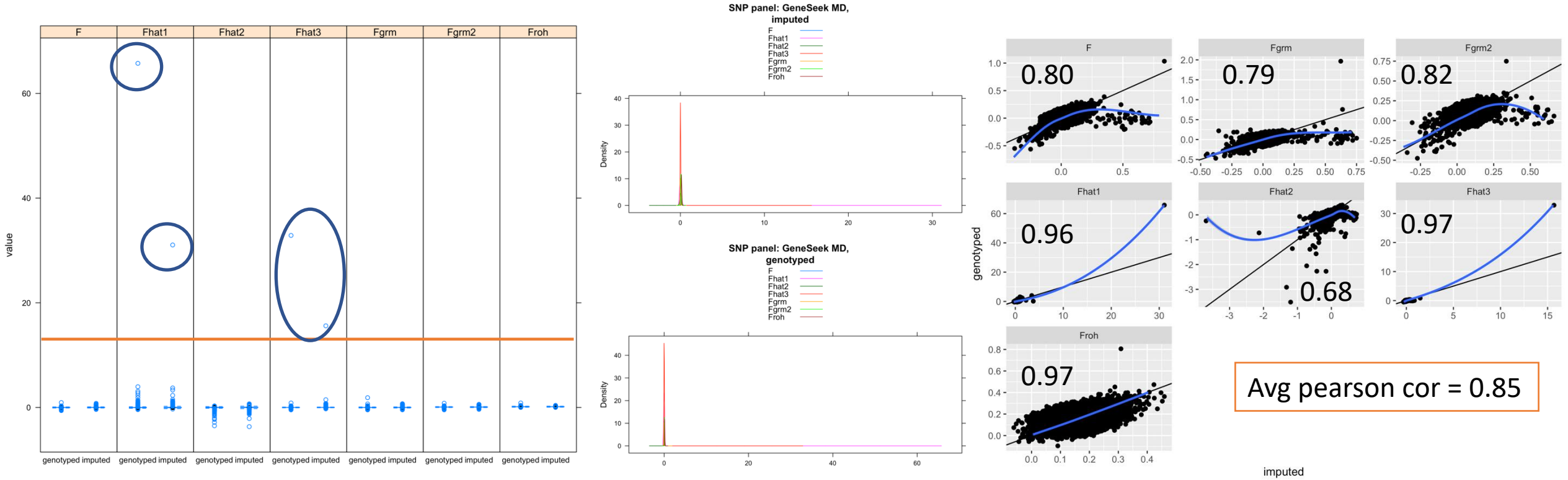
GeneSeek MD (n=12,030, nSNP=27,331 – MD)

genotyped

imputed

- Corr to Fped higher in genotyped, but fairly consistent between genotyped-imputed

Results and Discussion
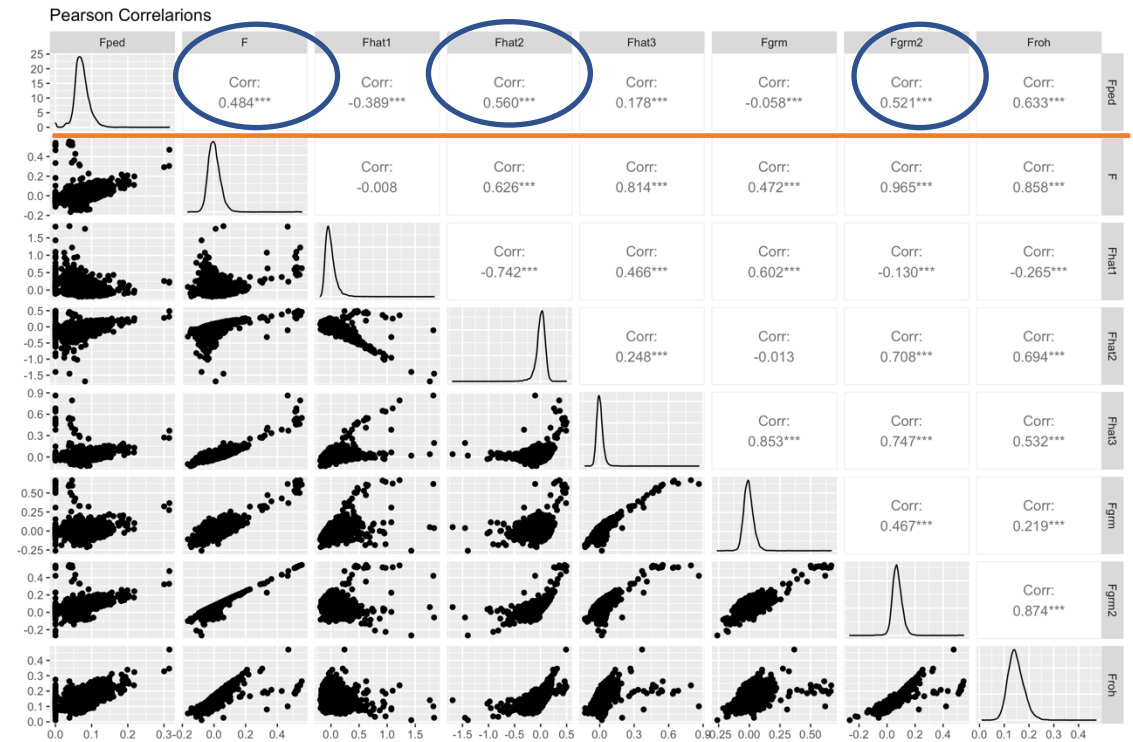
GeneSeek MD (n=12,030, nSNP=27,331 – MD)
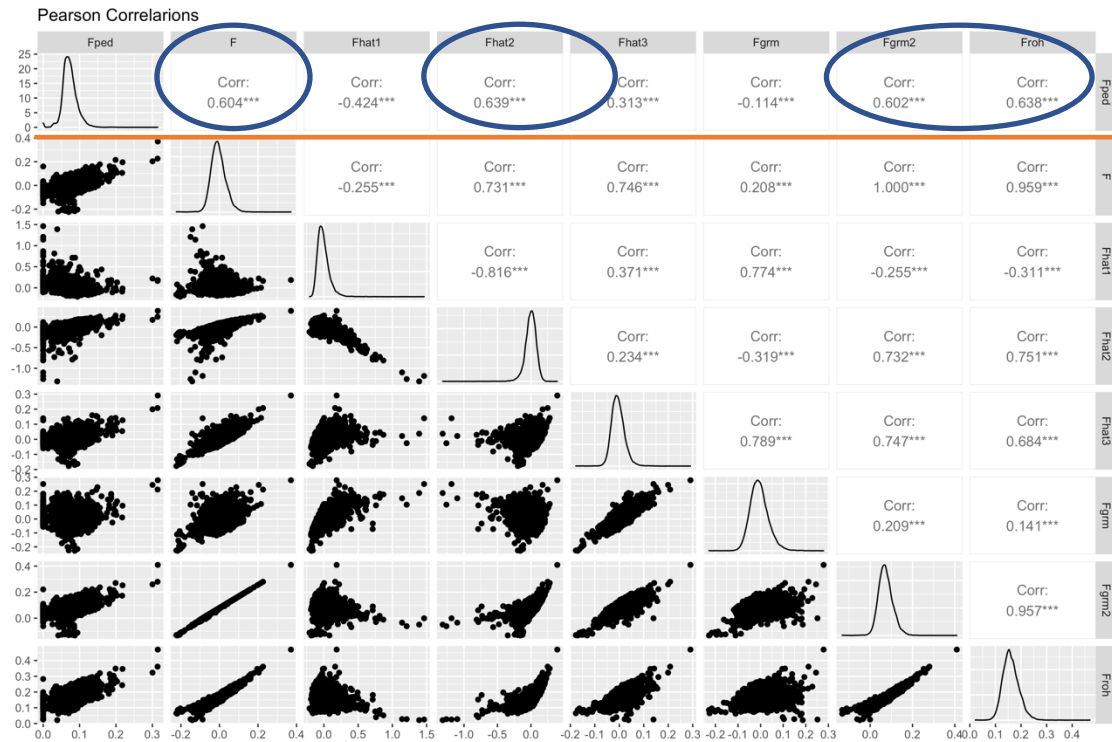

- Extreme values with genotyped and imputed

Results and Discussion

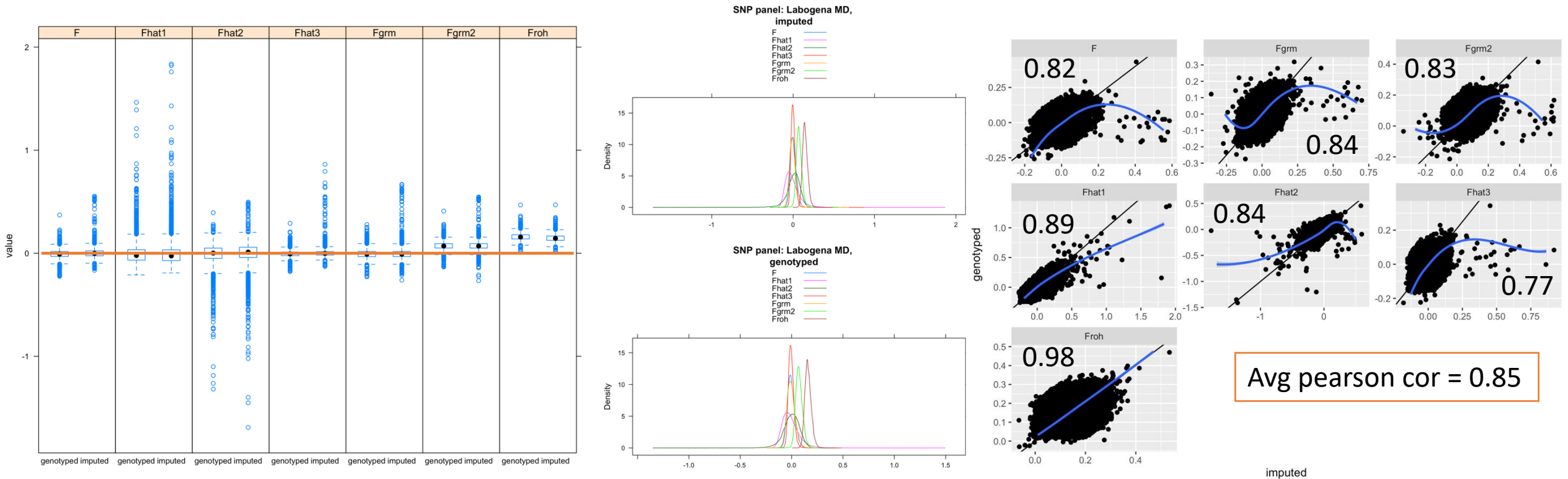**Labogena MD (n=10,705, nSNP=40,218 – MD)**

genotyped

imputed

- Corr to Fped higher in genotyped

# Results and Discussion

Labogena MD (n=10,705, nSNP=40,218 – MD)

- The imputed had always higher variability than the genotyped

Results and Discussion

Summary over SNP-panels across methods

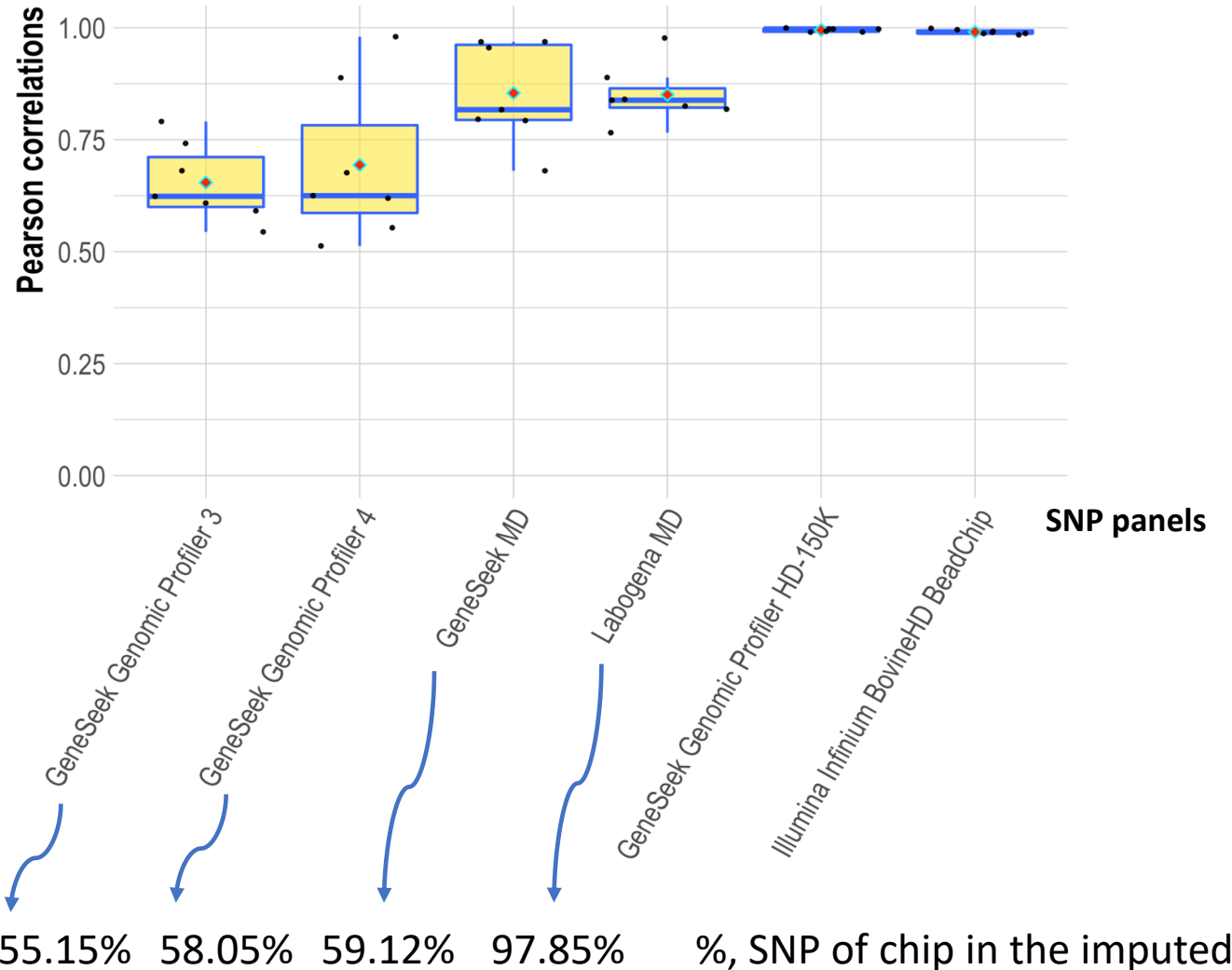Correlations of Genomic Inbreeding Coefficients
Imputed vs. Genotyped

Summary over methods across SNP-panels

Correlations of Genomic Inbreeding Coefficients
Imputed vs. Genotyped
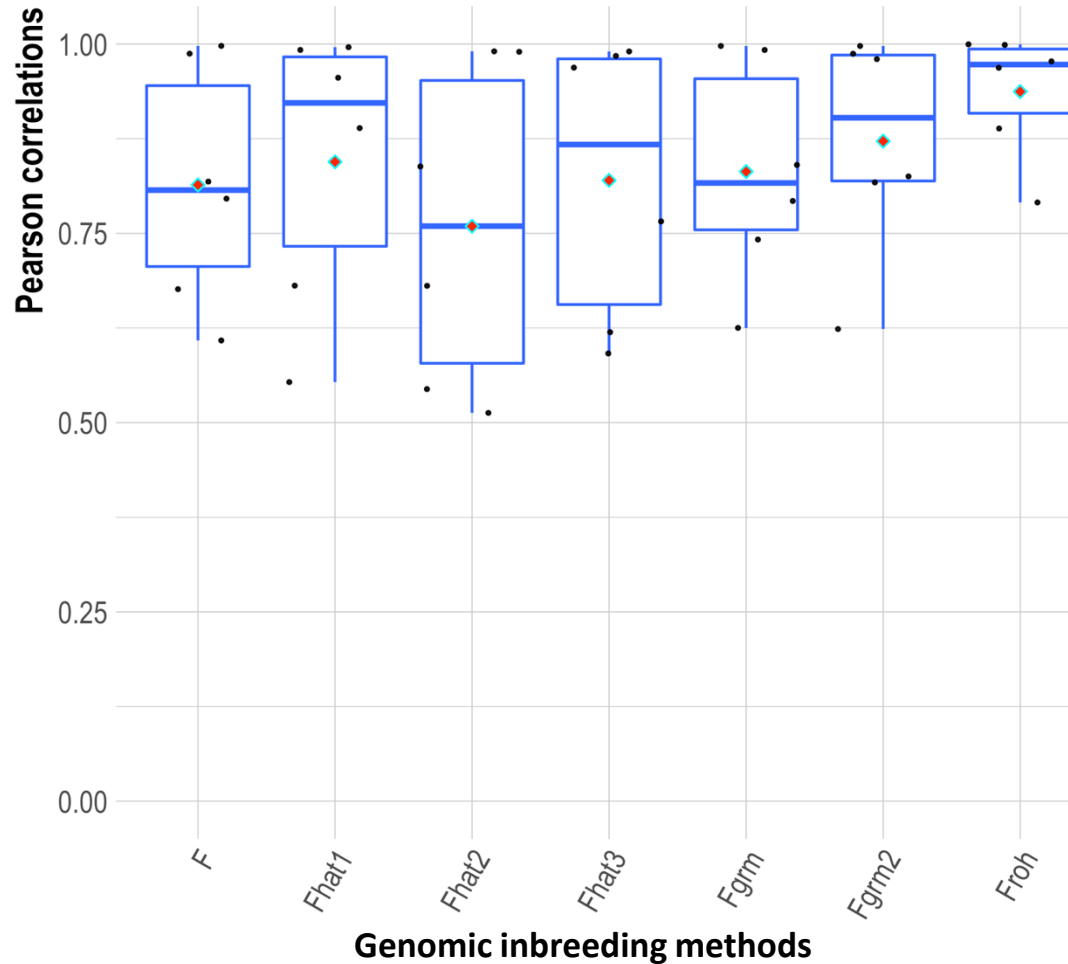
- FROH, more consistent across methods

55.15%   58.05%   59.12%   97.85%   %, SNP of chip in the imputed SNP

Conclusion

- ❑ FROH
  - Seems more robust
  - The only method with Fgen>0
  - Higher mean values
  - Always higher correlations with Fped
  - Generally lower variability
- ❑ HD
  - Imputed ≅ genotyped
  - Slightly better results of GeneSeek 150k vs. Illumina 777k
- ❑ MD
  - Correlations are, in general, driven by subgroups
  - The corFped,Fgen was higher for the genotyped vs. the imputed
  - Differences among SNP panels exist
  - Lagogena MD appears more robust compared to the rest MD tested
- ❑ Fhat2 ≠ F
- ❑ High extremes → GeneSeek 4 and GeneSeek MD!

Acknowledgements

Prof. Summer A.
Prof. Cipolat-Gotet C.
Prof. Cassandro M.
Dr. Marusi M.
Dr. Finocchiaro R.

Thank you for
your atttention!

ASPA 2021 PADOVA

24th CONGRESS
OF THE ANIMAL SCIENCE AND PRODUCTION ASSOCIATION
ANIMAL SCIENCE AND SOCIETY CONCERNS
SEPTEMBER 21-24, 2021 PADOVA (ITALY)

ANAFIBJ

UNIVERSITÀ DI PARMA