Genomic selection strategies and their potential to maintain rare alleles and de-novo mutations

M. Schrauf, J. Vandenplas, H. Mulder & Y. Wientjes

EAAP 2023 – Lyon, France
28th August 2023

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 101000226
Motivation

Rare alleles and de-novo mutations have...

- low correlation with phenotypes at the population level
- usually weak linkage with SNP markers
Rare alleles and de-novo mutations have...

- low correlation with phenotypes at the population level
- usually weak linkage with SNP markers

Led to think that

- genomic selection may not use favorable rare alleles effectively
- could loose rare alleles at a higher rate than pedigree selection
Previous works

Compared mass selection, pedigree selection and genomic selection

Compared mass selection, pedigree selection and genomic selection

Some conclusions about genomic selection:

- inclusion of own phenotypes is a main factor in the conservation of rare alleles
Compared mass selection, pedigree selection and genomic selection

Some conclusions about genomic selection:

- inclusion of own phenotypes is a main factor in the conservation of rare alleles
- doesn’t have to be worse than pedigree selection at this
Compared mass selection, pedigree selection and genomic selection

Some conclusions about genomic selection:

- inclusion of own phenotypes is a main factor in the conservation of rare alleles
- doesn’t have to be worse than pedigree selection at this
- but is much more prone, specifically, to hitch-hiking than pedigree selection

Assessment of different genomic selection strategies

Not *if* genomic selection but *how* genomic selection may be implemented
Assessment of different genomic selection strategies

Not *if* genomic selection but *how* genomic selection may be implemented

**Selection Strategies:**

**Truncation selection (TS) ------→** Maximize average EBVs from selected candidates

Optimal contributions (OCS) ----→ with a constraint on the candidates' coancestry
Meuwissen et al., (2020) Frontiers

Alleles re-weighting (ARW) ------→ with favorable rare alleles up-weighted in EBVs
Liu et al., (2015) GSE
(2 versions: *fixed* and *moving* time horizon)

Constrained allele loss (CAL) ------→ with a constraint on the reduction in frequency
novel strategy
of rare favourable alleles

*plus Random selection (RS) for reference*
Assessment of different genomic selection strategies

Not *if* genomic selection but *how* genomic selection may be implemented

Selection Strategies:

- **Truncation selection (TS) →** Maximize average EBVs from selected candidates
- **Optimal contributions (OCS) →** with a constraint on the candidates' coancestry
  - Meuwissen et al., (2020) Frontiers
- **Alleles re-weighting (ARW) →** with favorable rare alleles up-weighted in EBVs
  - Liu et al., (2015) GSE
  - (2 versions: *fixed* and *moving* time horizon)
- **Constrained allele loss (CAL) →** with a constraint on the reduction in frequency of rare favourable alleles
  - novel strategy

*plus Random selection (RS) for reference*
Introduction

Current work

- Assessment of different genomic selection strategies

Not *if* genomic selection but *how* genomic selection may be implemented

Selection Strategies:

Truncation selection (TS) \(\rightarrow\) Maximize average EBVs from selected candidates

Optimal contributions (OCS) \(\rightarrow\) with a constraint on the candidates' coancestry

Meuwissen et al., (2020) Frontiers

Alleles re-weighting (ARW) \(\rightarrow\) with favorable rare alleles up-weighted in EBVs

Liu et al., (2015) GSE

(2 versions: *fixed* and *moving* time horizon)

Constrained allele loss (CAL) \(\rightarrow\) with a constraint on the reduction in frequency of rare favourable alleles

*plus Random selection (RS) for reference*
Assessment of different genomic selection strategies

Not *if* genomic selection but *how* genomic selection may be implemented

Selection Strategies:

- **Truncation selection (TS)** $\rightarrow$ Maximize average EBVs from selected candidates
- **Optimal contributions (OCS)** $\rightarrow$ with a constraint on the candidates' coancestry
  - Meuwissen et al., (2020) Frontiers
- **Alleles re-weighting (ARW)** $\rightarrow$ with favorable rare alleles up-weighted in EBVs
  - Liu et al., (2015) GSE
  - (2 versions: *fixed* and *moving* time horizon)
- **Constrained allele loss (CAL)** $\rightarrow$ with a constraint on the reduction in frequency of rare favourable alleles
  - novel strategy

*plus Random selection (RS) for reference*
Assessment of different genomic selection strategies

Not *if* genomic selection but *how* genomic selection may be implemented

**Selection Strategies:**

- **Truncation selection (TS) →** Maximize average EBVs from selected candidates
- **Optimal contributions (OCS) →** with a constraint on the candidates' coancestry
  - Meuwissen et al., (2020) Frontiers
- **Alleles re-weighting (ARW) →** with favorable rare alleles up-weighted in EBVs
  - Liu et al., (2015) GSE
  - (2 versions: *fixed* and *moving* time horizon)
- **Constrained allele loss (CAL) →** with a constraint on the reduction in frequency of rare favourable alleles
  - novel strategy

*plus Random selection (RS) for reference*
The Simulation

The Population:
50 discrete generations
1000 individuals
100 sires + 100 dams selected
- selected without own phenotypes
- using marker effects learnt from the 3 prior generations

Genome:
20k SNP marker panel
- MAFs 0.5 to 0.1
- neutral loci
2k starting causal loci
mutations rate $3.8 \times 10^{-5} \text{ (loci.ind)}^{-1}$

Simulation approach from Wientjes, et al. 2022
## The Simulation

<table>
<thead>
<tr>
<th>The Population:</th>
<th>Genome:</th>
</tr>
</thead>
<tbody>
<tr>
<td>50 discrete generations</td>
<td>20k SNP marker panel</td>
</tr>
<tr>
<td>1000 individuals</td>
<td>- MAFs 0.5 to 0.1</td>
</tr>
<tr>
<td>100 sires + 100 dams selected</td>
<td>- neutral loci</td>
</tr>
<tr>
<td>- selected without own phenotypes</td>
<td>2k starting causal loci</td>
</tr>
<tr>
<td>- using marker effects learnt from the 3 prior generations</td>
<td>mutations rate $3.8 \times 10^{-5} (\text{loci.ind})^{-1}$</td>
</tr>
</tbody>
</table>

Simulation approach from Wientjes, et al. 2022
The Traits

Additive
Normally distributed additive effects, with a common variance.

Dominant
Includes dominance effects, with a small positive bias for heterozygotes.

Epistatic
Includes pairwise interactions, with connectivity pattern taken from a yeast study.

Traits specifications taken from Wientjes, et al. 2022

Yeast study in Costanzo et al., 2016
Results & Discussion
Results & Discussion

Evolution of total genetic value

<table>
<thead>
<tr>
<th></th>
<th>Additive</th>
<th>Epistatic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Generation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total genetic value</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Strategy:
- RS
- TS
- OCS
- ARWF
- ARWm
- CAL
Results & Discussion

Evolution of additive genetic variance

<table>
<thead>
<tr>
<th>Generation</th>
<th>Additive</th>
<th>Epistatic</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>35</td>
<td></td>
<td></td>
</tr>
<tr>
<td>40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>45</td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Additive genetic variance

- Strategy:
  - RS
  - TS
  - OCS
  - ARWF
  - ARWm
  - CAL

Graph showing the evolution of additive genetic variance over generations for both additive and epistatic effects.
Genetic gain vs. genetic variance

Alternative strategies compared with truncation selection (Additive)
Results & Discussion

Genetic gain vs. genetic variance

Alternative strategies compared with truncation selection (Additive)
Results & Discussion

Genetic gain vs. genetic variance

Alternative strategies compared with truncation selection (Additive)
Results & Discussion

Genetic gain vs. genetic variance

Alternative strategies compared with truncation selection (Additive)

ARW strategies allow effective trade-off between increased genetic gain and conservation of genetic variance.
Considering traits with non-additive effects improves the assessments of OCS and ARWm for genetic gain.
Selection of de-novo mutations
Results

Selection of de-novo mutations

Contribution of DNMs to TBVs

Number of DNMs (Additive trait)

Strategy
- RS
- TS
- OCS
- ARWf
- ARWm
- CAL

Mutation type
- favorable
- deleterious
Results

Selection of de-novo mutations

No strategy outperforms truncation selection on these metrics

All selection strategies are applying pressure on the mutations
Considering traits with non-additive effects, selection of DNMs becomes more challenging.

CAL selection has the lowest and OCS the highest contribution of DNMs to TBVs.
For the fully additive trait

- Truncation selection starts with higher gains,
  - Saturates earlier and gain is surpassed by a reweighting strategy.

- Allelic reweighting is an effective strategy for long term selection,
  - Even if working with markers rather than causal loci.

- No strategy is significantly more effective at keeping favourable de-novo mutations segregating,
  - Although they are all slowly purging the deleterious mutational load.
For the trait with epistasis

- Allelic reweighting remains an effective strategy for long term selection,
  - Even while favorable alleles change through generations.

- Optimal contribution outperforms truncation’s long term genetic gain,
  - Which didn’t happen for the fully additive trait.

- Purging deleterious mutations becomes more challenging for all the selection criteria explored,
  - Possibly due to a combination of lower narrow-sense heritability and changes in which rare alleles are estimated to be favorable.
Thank you for your attention

www.rumigen.eu
Maximizes average EBVs from selected candidates without any consideration of diversity management.

We estimated SNP effects ($\beta$) with the phenotypes of the 3 previous generations (by means of a SNPBLUP model).

And selected the 100 top sires and 100 top dams for:

$$\text{GEBVs} = X\beta$$
Optimal contribution selection (OCS)

Maximize average EBVs from selected candidates with a constraint on the candidates' coancestry

Maximize $\mathbf{g} = \mathbf{c}'\mathbf{X}\mathbf{\beta}$

$K_t \geq \frac{1}{2} \mathbf{c}'\mathbf{G}\mathbf{c}$

$Q_c = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \end{bmatrix}'$

$c \geq 0$

where $K_t = K_{t-1} + (1-K_{t-1})/(2Ne)$, using $Ne=60$

Marker effects of rare alleles re-weighted according to Liu et al., 2015

\[ W_{jj} \propto 1/p_j^{c(t)} \]

where \( c(0) = 0.5 \) and \( c(T) = 0.0 \) and \( p_j \) is the freq of the favourable allele.

\[ w \text{GEBVs} = Xw\beta \]

(years to horizon; dotted line: 5 years, solid line: 20 years)
Included two variants of this strategy, using different definitions for the time horizons:

- **ARWf (fixed)**: using the full length of the simulation of 50 generations, as the time horizon.
- **ARWm (moving)**: using a moving horizon, always 5 generations ahead.
Maximize average EBVs from selected candidates with a constraint on the loss of rare (favourable) alleles.

Maximize \( g = c'X\beta \)

\[ L \geq c'X\alpha \]

\[ Qc = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \end{bmatrix} \]

\( c \geq 0 \)

where \( \alpha_j = -\log(1/n \ast (1 + (J'X)_j)) \) [if \( \beta_j \geq 0 \)],

\( L = 1.10*1/n*(J'X\alpha) \), and \( J \) is an n-length vector of ones.
Results

Genetic gain vs. genetic variance

Genetic improvement vs. reduction in genetic variance

ARW strategies allow effective trade-off between increased genetic gain and conservation of genetic variance.