

AlphaPart - R implementation of the method for partitioning genetic trends

Jana Obšteter¹, Justin Holl², John M. Hickey³, and Gregor Gorjanc³

¹Department of Animal Science, Agricultural Institute of Slovenia, Hacquetova ulica 17, 1000 Ljubljana, Slovenia

²The Pig Improvement Company, Genus plc, 100 Bluegrass Commons Blvd., Ste 2200, Hendersonville, TN 37075, USA

³The Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh, Easter Bush, Midlothian, EH259RG, United Kingdom

*Jana Obšteter, Agricultural Institute of Slovenia, Department of Animal Science, Hacquetova ulica 17, 1000 Ljubljana, Slovenia

E-mail addresses:

JO: jana.obsteter@kis.si

JH: justin.holl@genusplc.com

JMH: john.hickey@roslin.ed.ac.uk

GG: gregor.gorjanc@roslin.ed.ac.uk

21 Abstract

22 **Background:** In this paper we present the AlphaPart R package, an open-source software that
23 implements a method for partitioning breeding values and genetic trends to identify sources of
24 genetic gain. Breeding programmes improve populations for a set of traits, which can be
25 measured with a genetic trend calculated from averaged year of birth estimated breeding
26 values of selection candidates. While sources of the overall genetic gain are generally known,
27 their realised contributions are hard to quantify in complex breeding programmes. The aim of
28 this paper is to present the AlphaPart R package and demonstrate it with a simulated ~~pig~~
29 ~~breeding example~~stylized multi-tier breeding example, usually applied in poultry and pig
30 breeding.

31 **Results:** The package includes the main partitioning function AlphaPart, that partitions the
32 breeding values and genetic trends by analyst defined paths, and a set of functions for
33 handling data and results. The package is freely available from CRAN repository at
34 <http://CRAN.R-project.org/package=AlphaPart>. We demonstrate the use of the package by
35 examining the genetic gain in a ~~pig~~stylized multi-tier breeding ~~example~~programme, in which
36 the multiplier achieved higher breeding values than the nucleus for traits measured and
37 selected in the multiplier. The partitioning analysis revealed that these higher values depended
38 on the accuracy and intensity of selection in the multiplier and the extent of gene flow from
39 the nucleus. For traits measured only in the nucleus, the multiplier achieved comparable or
40 smaller genetic gain than the nucleus depending on the amount of gene flow.

41 **Conclusions:** AlphaPart implements a method for partitioning breeding values and genetic
42 trends and provides a useful tool for quantifying the sources of genetic gain in breeding

programmes. The use of AlphaPart will help breeders to better understand or improve their breeding programmes.

Keywords: genetic trend, partition, Mendelian sampling term, R package, pigmulti-tier breeding

Background

In this paper we present the AlphaPart R package that implements a method for partitioning breeding values and genetic trends, ~~We and demonstrate~~ the package with a ~~pig~~stylized multi-tier breeding example usually applied in poultry and pig breeding. Breeding programmes improve populations for a set of traits by selecting and ~~inter~~mating genetically superior individuals. Population improvement can be measured with a genetic trend calculated from averaged year of birth estimated breeding values of selection candidates [1,2].

Partitioning method

While sources of the overall genetic gain are generally known, their realised contributions are hard to quantify in complex breeding programmes. García-Cortés *et al.* [3] proposed a method for such analysis. ~~For completeness of the manuscript we briefly summarize the method below. First, the method partitions breeding values into parent average and Mendelian sampling terms [4], and allocates the terms to analyst-defined “paths” (males, females, tested sires, etc.). Next, it summarizes path specific terms to quantify path contributions to the overall genetic trend.~~

The infinitesimal model assumes that the breeding values follow a normal distribution with mean zero and variance $\mathbf{A}\sigma_a^2$, where \mathbf{A} is the pedigree relationship matrix and σ_a^2 is the genetic variance for the trait in the founder population. We can partition non-founders' breeding value a into parent average ($0.5a_{sire} + 0.5a_{dam}$) and Mendelian sampling term (w) [4], which arises from independent segregation and recombination of parental chromosomes. For founders we have no parental information, hence we assign their breeding value completely to Mendelian sampling term.

We can decompose the variance of the breeding values to $\mathbf{A}\sigma_a^2 = \mathbf{TDT}'\sigma_a^2$, where \mathbf{T} is a triangular matrix relating individuals to their ancestors (it holds information on gene transmission), and $\mathbf{D}\sigma_a^2$ includes the variance of the breeding values of the founders and variance of the Mendelian sampling terms of the non-founders. Following this, we can express additive genetic values $\mathbf{a} = \mathbf{T}\mathbf{w}$, where \mathbf{w} follows a normal distribution with mean zero and variance $\mathbf{D}\sigma_a^2$. Hence, we can further partition an individual's breeding value into the contribution of founders breeding values and non-founders' Mendelian sampling terms (including the individual itself) as

$$a_i = \sum_{j=1}^i T_{ij} w \quad [\text{Woolliams et al., 1999}]$$

We can then specify different groups generating the Mendelian sampling term and assign individuals to the groups (called "path" in the Implementation). If we summarise the group-specific contributions, we can partition the breeding values into $\mathbf{a} = \mathbf{a}_1 + \mathbf{a}_2 + \dots + \mathbf{a}_m$, where 1 to m are different groups, and quantify group contributions to the overall genetic trend.

The partitioning method has been used in a number of cases. Gorjanc *et al.* [5] and Gorjanc *et al.* [6] estimated contributions of national breeding programmes to Brown-Swiss and Holstein country-specific and global genetic trends. Špehar *et al.* [7] estimated contributions of national selection and importation in Croatian Simmental cattle. Škorput *et al.* [8] estimated the contribution of national selection and importation in two pig breeds in Croatia, and extended the analysis with the quantification of uncertainty [2]. However, these studies used ~~manual dedicated software~~ implementations of the partitioning method, for which no open-source software exists.

The aim of this paper is to: i) present the AlphaPart R package; and ii) demonstrate it with a simulated ~~pigstylized multi-tier~~ breeding example that quantifies nucleus-multiplier gene flow and the contribution of nucleus and multiplier selection on genetic gain in the two tiers.

Implementation

We first demonstrate the AlphaPart package and its functions on an example dataset. Next, we describe the simulation of a ~~pigstylized multi-tier~~ breeding ~~exampleprogramme~~ to demonstrate the use of AlphaPart.

AlphaPart

AlphaPart is an R package available from CRAN repository at <https://CRAN.R-project.org/package=AlphaPart>. It consists of the main function `AlphaPart()` for partitioning breeding values and auxiliary functions for manipulating data and summarizing, visualizing, and saving results. ~~The package includes an example dataset AlphaPart.ped, which includes a four-generation pedigree and information about~~

105 | ~~the generation, country, gender, and breeding values. Below we describe and demonstrate the~~
106 | ~~functions with the dataset.~~

107 | We install and load the package with:

```
108 > install.packages(pkg = "AlphaPart")
```

```
109 > library(package = "AlphaPart")
```

110 | The package includes an example dataset `AlphaPart.ped`. We show the first few lines
111 | below. ~~which~~ The dataset includes a four-generation pedigree (individual `IIId`, father `FIId`,
112 | and mother `MIId`) and information about the generation (`gen`), country, gender, and breeding
113 | values for trait 1 (`bv1`) and trait 2 (`bv2`). ~~Below we describe and demonstrate the functions~~
114 | ~~with the dataset.~~

	<code>IIId</code>	<code>FIId</code>	<code>MIId</code>	<code>gen</code>	country	gender	<code>bv1</code>	<code>bv2</code>
A				1	domestic	F	100	88
B				1	import	M	105	110
C	B	A		2	domestic	F	104	100

115

116 | Here we describe and demonstrate the functions with the `AlphaPart.ped` dataset. We use
117 | the `AlphaPart()` function to partition breeding values (`bv1`) in the `AlphaPart.ped` by
118 | the country variable into domestic and import contributions:

```
119 > data(AlphaPart.ped)
```

```
120 > part <- AlphaPart(x = AlphaPart.ped,
```

```
121                       colPath = "country",
```

```
122                       colBV = "bv1")
```

The output part is an AlphaPart object, which is a list with an info element and an element for each of the partitioned traits. The info element is a list with information on the path variable (path), number of path levels (nP), names of path levels (lP), number of traits (nT), names of traits (lT), and putative warnings (warn). The trait element (shown below) bv1 is a table. It includes the input information, the breeding value for the trait (bv1), the parent average (bv1_pa), the Mendelian sampling term (bv1_w), and the contributions of each level of the path variable to the breeding value (bv1_domestic, bv1_import).

```
> part$bv1
```

IIId	FId	MId	gen	country	gender	bv1	bv1_pa	bv1_w	bv1_domestic	bv1_import
A			1	domestic	F	100	0	100	100	0
B			1	import	M	105	0	105	0	105
C	B	A	2	domestic	F	104	102.5	1.5	51.5	52.5
T	B		2	import	F	102	52.5	49.5	0	102

The partitioning function AlphaPart() requires a data frame holding pedigree with animal/sire/dam or animal/sire/maternal-grandsire, a time-ordering variable such as year of birth, partition variable (path), and breeding values. Following the method described in García-Cortés *et al.* [3], [wethe function](#) recurses the pedigree from the oldest to the youngest individual, for each individual calculate parent average and Mendelian sampling terms for any number of traits and assign terms to paths. We partition multiple traits by specifying a vector of variables, say colBV = c("bv1", "bv2"). The multiple trait option can also serve to partition samples from a posterior distribution [of breeding values](#) to quantify uncertainty [2, 8]. To speed-up calculations we use C++ and trait-vectorised partitioning. The function can

141 also directlysimultaneously partition and summarize path contributions by a grouping
 142 variable“on-the-fly”, which is a useful computational speed-up for huge pedigrees.
 143 Alternatively, we subsequently use summary() function to summarize the partitions. The
 144 output object of the AlphaPart() function is either AlphaPart or
 145 summaryAlphaPart class.

146 We use the generic summary.AlphaPart() function to summarize an AlphaPart
 147 object by a grouping variable, say generation (gen):

```
148 > sumPartByGen <- summary(part, by = "gen")
```

149 The output is an summaryAlphaPart object. It is a list with an info element and (as in
 150 AlphaPart object) and a bv1 element showed below, that we can display with print()
 151 function. The N columns shows the number of individuals in each generation. We used the
 152 default mean function to summarize the data, so Sum columns shown the mean of the
 153 breeding values of individuals in each generation. The columns domestic and import show the
 154 mean contribution of domestic selection and import.

```
155 > print(sumPartByGen)
```

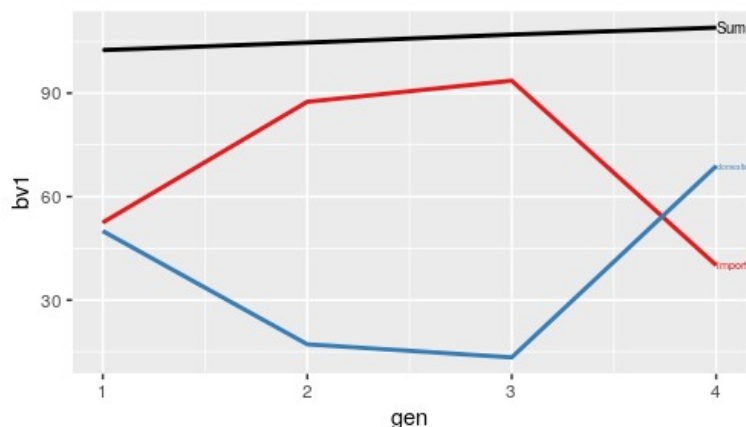
<u>gen</u>	<u>N</u>	<u>Sum</u>	<u>domestic</u>	<u>import</u>
<u>1</u>	<u>2</u>	<u>102.5</u>	<u>50</u>	<u>52.5</u>
<u>2</u>	<u>3</u>	<u>104.6667</u>	<u>17.16667</u>	<u>87.5</u>
<u>3</u>	<u>2</u>	<u>107</u>	<u>13.375</u>	<u>93.625</u>
<u>4</u>	<u>1</u>	<u>109</u>	<u>68.875</u>	<u>40.125</u>

156 The summary() function summarizes breeding values and their path partitions by levels of
 157 grouping variable. By default, we summarize with a mean, but the user can specify any R

158 function via the `FUN` argument. The `summary()` function can also summarize only a subset
 159 of the object via the `subset` argument.

160 We use the generic `plot.summaryAlphaPart` function to plot summarized partitions:

161 `> plot(sumPartByGen)`



162

163 We also provide a number of utility functions that ease partitioning analysis. ~~With the~~
 164 `pedFixBirthYear()` function ~~we impute~~imputes missing or fix erroneous years of birth.
 165 ~~The~~With the `pedSetBase()` function ~~sets~~we set the base population by specifying founders
 166 and removing older pedigree records. ~~The~~With the `AlphaPartSubset()` function ~~we~~
 167 keeps partitions for specified paths in the `AlphaPart` or `summaryAlphaPart` objects.
 168 ~~I~~With the `AlphaPartSum()` function ~~we sum~~sums the partitions of several paths in a
 169 `summaryAlphaPart` object. The `AlphaPartSubset()` and `AlphaPartSum()`
 170 functions simplify the presentation of partitioning analysis.

<u>Function</u>	<u>Description</u>
<u><code>pedFixBirthYear()</code></u>	<u>Imputes missing or fixes erroneous years of birth.</u>
<u><code>pedSetBase()</code></u>	<u>Sets the base population by specifying founders and</u>

	<u>removing older pedigree records.</u>
<u>AlphaPartSubset()</u>	<u>keeps partitions for specified paths in the AlphaPart or summaryAlphaPart objects.</u>
<u>AlphaPartSum()</u>	<u>Sums the partitions of several paths in a summaryAlphaPart object.</u>

171

172 | Stylized multi-tier breeding~~Pig breeding~~ example

173 | We applied the AlphaPart R package ~~onto~~ a simulated ~~pig~~stylized multi-tier breeding
174 | example, usually applied in poultry and pig breeding. ~~To~~We aimed to examine the nucleus-
175 | multiplier gene flow and the contribution of nucleus and multiplier selection on genetic gain
176 | in ~~both~~each tiers. ~~B~~Pig breeders select in the nucleus and multiply this improvement in the
177 | multiplier to supply a large number of commercial animals. The multiplier generally has
178 | lower genetic mean than the nucleus due to time-lag. However, animals with very high
179 | breeding values are often observed in the multiplier for some traits and we aimed to use
180 | AlphaPart to explain the source of this observation. To this end we have first simulated a
181 | stylised ~~pig~~multi-tier breeding programme that exposes the drivers of real observations. We
182 | have next partitioned the genetic trend of true breeding values by a tier-gender variable to
183 | quantify sources of genetic gain in the nucleus and the multiplier.

184 | We used the AlphaSimR package [9] to simulate a ~~pig~~multi-tier breeding programme for a
185 | single breed with 40 years of selection on two uncorrelated traits. Trait 1 had heritability 0.25
186 | and trait 2 had heritability 0.10. We measured both traits in the nucleus and only trait 1 in the
187 | multiplier. We selected on the index of the two traits with equal emphasis. We split the
188 | simulation into initial 20 years of a burn-in and 20 years of evaluation.

189 In the burn-in we simulated only the nucleus and selected animals based on the index of
 190 phenotype values for both traits. We selected 25 males and 500 females each year and
 191 randomly crossed them to produce a new generation of 6,000 progeny (12 per cross). At the
 192 end of the burn-in we generated 5,000 females to seed the multiplier.

193 In the evaluation we simulated both the nucleus and the multiplier and selected animals based
 194 on the index of estimated breeding values for both traits. In the nucleus, we selected 25 males
 195 and 500 females each year and randomly crossed them to produce a new generation of 6,000
 196 progeny (12 per cross). In the multiplier, we selected 750 females each year and randomly
 197 crossed them to a set of males to produce a new generation of 9,000 progeny (12 per cross).
 198 To quantify the effect of selection in the multiplier on genetic gain we defined the set of males
 199 as either 1) the 25 best nucleus males (MaleFlow100 scenario) or 2) the 25 best nucleus males
 200 and 100 best multiplier males (MaleFlow20 scenario).

201 We estimated the breeding values for each trait independently before each nucleus or
 202 multiplier selection decision. We ran [a](#) pedigree-based model implemented in blupf90 [10]
 203 and used all available data from evaluation years. The model included the mean as a fixed
 204 effect and animal breeding values as a random effect modelled hierarchically with pedigree.

205 Finally, we partitioned the true breeding values and genetic trends with the [AlphaPart](#) as
 206 demonstrated above. We used `AlphaPart()` function to partition [standardized](#) true
 207 breeding values from the 20 evaluation years by the tier-gender variable and
 208 `summary.AlphaPart()` function to summarize the partitions by year to quantify the
 209 contribution of each tier-gender level to genetic trend in the nucleus and the multiplier.

210

[Generation](#) [IIId](#) [FIId](#) [MIId](#) [TierGender](#) [TbvT1_s](#) [TbvT2_s](#) [TbvI_s](#)

<u>20</u>	<u>111670</u>	<u>107396</u>	<u>106535</u>	<u>GN-M</u>	<u>-0.17</u>	<u>0.08</u>	<u>-0.04</u>
<u>20</u>	<u>111671</u>	<u>107396</u>	<u>106535</u>	<u>GN-F</u>	<u>-0.55</u>	<u>0.05</u>	<u>-0.25</u>
<u>20</u>	<u>111672</u>	<u>107396</u>	<u>106535</u>	<u>GN-M</u>	<u>-0.31</u>	<u>-0.17</u>	<u>-0.24</u>
<u>20</u>	<u>111673</u>	<u>107396</u>	<u>106535</u>	<u>GN-F</u>	<u>0.70</u>	<u>-0.68</u>	<u>0.01</u>

```
# Partition the standardized breeding value for trait 1, trait 2 and index
Part1g <- AlphaPart(x = as.data.frame(PedEval1), sort = FALSE,
                    colId = "IId", colFid = "FId", colMid = "Mid",
                    colPath = "TierGender",
                    colBV = c("TbvT1_s", "TbvT2_s", "TbvI_s"))
```

We repeated the simulation 10 times. We present standardized true breeding values and genetic trends, as well as their partitions with mean set to zero and genetic standard deviation set to one in the year 20. We chose to present true (instead of estimated) breeding values to assess the true sources of genetic gain. The simulation code for the datasets generated and/or analysed during the current study are available in the GitLab repository, https://git.ecdf.ed.ac.uk/HighlanderLab_public/jobsteter_alphapart.

Results

The results show partitions of true breeding values and genetic trends obtained with the AlphaPart for the two simulated **pigstylized multi-tier** breeding scenarios. Partitioning showed that we can explain the situation with very high breeding values in the multiplier by the extent of nucleus-multiplier gene flow as well as accuracy and intensity of multiplier selection.

Partitioning the true breeding values and genetic trend of MaleFlow100 scenario

230 In MaleFlow100 scenario the multiplier achieved a higher final genetic gain than the nucleus
231 for trait 1 due to selection of multiplier females. We show this in Figure 1 that presents the
232 distribution of true breeding values and their partitions by trait and tier for two years of one
233 replicate of MaleFlow100 scenario, and in Figure 2 that presents the genetic trends and their
234 partitions summarised over 10 replicates. As expected, ~~The partitioning~~ expectedly showed
235 that in the nucleus the genetic gain stemmed from selection of nucleus males and nucleus
236 females. However, the contribution of male and female selection changed over the years.
237 While in year 23 the contributions of male and female selection were more comparable, by
238 year 40 male selection contributed more. The mean final genetic gain in the nucleus for trait 1
239 was 9.75 and 8.34 for trait 2, with male selection contributing 5.65 for trait 1 and 4.92 for trait
240 2, and female selection contributing 4.10 for trait 1 and 3.42 for trait 2.

241 In the multiplier, the genetic gain was higher than in the nucleus. In year 23 the multiplier had
242 higher genetic gain than the nucleus for both traits, while in year 40 it had higher genetic gain
243 only for trait 1. The higher genetic gain was partly due to larger contribution of nucleus males
244 in multiplier than in nucleus (via gene flow) and partly due to non-zero contribution of
245 multiplier female selection. The mean final genetic gain in the multiplier for trait 1 was 10.00
246 with nucleus males contributing 5.75, nucleus females 4.09, and multiplier females 0.14. The
247 mean final genetic gain and path partitions for trait 2 in the multiplier were comparable to the
248 nucleus.

249 **Partitioning the true breeding values and genetic trend of MaleFlow20 scenario**

250 In the MaleFlow20 scenario selection of multiplier males further increased the final genetic
251 gain for trait 1 in the multiplier compared to the nucleus, but decreased the final genetic gain
252 for trait 2. We show this in Figure 3 that presents the distribution of true breeding values and

their partitions by trait and tier for two year of one replicate of MaleFlow20 scenario, and Figure 4 that presents the genetic trends and their partitions summarised over 10 replicates. As in MaleFlow100 scenario, the nucleus genetic gain stemmed from selection of nucleus males and females. Progressing from year 23 to year 40, the contribution of nucleus males increased compared to nucleus females, but only for trait 1. The mean final genetic gain for trait 1 was 10.09 and 8.39 for trait 2, with nucleus males contributing 5.69 for trait 1 and 5.17 for trait 2, and nucleus females contributing 4.40 for trait 1 and 3.22 for trait 2.

In the multiplier the genetic gain was again higher than in the nucleus, but only for trait 1. This higher genetic gain was consistent throughout the generations and was a result of selection of multiplier females and multiplier males, and reduced contribution of gene flow from the nucleus females (via reduced use of nucleus males). For trait 2, the multiplier performed progressively worse relative to the nucleus over the years. In year 40, the genetic gain in the multiplier was lower than in the nucleus due to a small average negative contribution of multiplier females and multiplier males and reduced contribution of gene flow from the nucleus females and nucleus males. The mean final genetic gain in the multiplier was 10.36 for trait 1 and 8.14 for trait 2, with nucleus males contributing 5.70 for trait 1 and 5.09 for trait 2, nucleus females contributing 4.21 for trait 1 and 3.13 for trait 2, multiplier males contributing 0.15 for trait 1 and -0.03 for trait 2, and multiplier females contributing 0.30 for trait 1 and -0.05 for trait 2.

Discussion

In this paper we present AlphaPart, freely available R package that implements the method for partitioning breeding values and genetic trends. We demonstrate the package on a simulated [pigstylized multi-tier](#) breeding example with a higher genetic trend for some traits in the

multiplier compared to the nucleus. Following this, we organized the discussion into two parts: i) advantages and disadvantages of the AlphaPart R package; ii) partitioning results of the pig-breeding example.

AlphaPart

AlphaPart is the first free implementation of the method for partitioning breeding values and genetic trends. The method and the package are valuable for deciphering and quantifying the sources of genetic gain in breeding programmes. The package is easy to use, since it streamlines the partitioning analysis into a few lines of R code. AlphaPart presents a holistic tool to perform a partitioning analysis, from preparing the input data - such as manipulating the pedigree data - to handling of results and plotting. The partitioning step is fast, even for large pedigrees, since the main partitioning function is recursive and implemented in C++.

AlphaPart is aimed at the researchers who are interested in quantifying the sources of genetic gain in their breeding programmes either to understand the dynamics of genetic gain, improve efficiency, assess the performance of different breeding actions, optimize investments etc. Such users should take into account the accuracy of the estimated breeding values and their Mendelian sampling terms, which are driven by the biology of the trait and breeding programme structure.

Our future work on AlphaPart will include extending the partitioning method in three areas. The first extension will utilise genomic information to inform which genome regions drive genetic change and what are sources of specific haplotypes or alleles. The second extension will use the partitioning method to analyse changes in genetic variance in addition to the genetic mean. The third extension will simplify handling of uncertainty of path contributions when working with samples from posterior distributions [2, 8].

Stylized multi-tier Pig breeding breeding example

The pig-multi-tier breeding example showed the investigative power of the partitioning method and the free AlphaPart implementation. Here we discuss the sources of genetic gain in the two tiers of a pig breeding programme.

By partitioning the genetic trend in a simulated pigmulti-tier breeding programme, we disentangled the observation of some multiplier animals having higher breeding values for some traits compared to the nucleus animals. While larger number of recombinations in the multiplier can potentially reveal more variation and occasional outlying animals, we expect lower breeding values in the multiplier due to time-lag between the nucleus and multiplier. The partitioning revealed that the gene flow from the nucleus into the multiplier was the main source of genetic gain in the multiplier, with the nucleus males contributing the most. This was expected due to nucleus-multiplier gene flow and higher intensity of selection in males.

However, the results also showed that selection in the multiplier can contribute genetic gain in addition to the gene flow from the nucleus. The multiplier outperformed the nucleus for trait 1, because with the 10,500 recorded multiplier animals there was substantial amount of information for accurate multiplier selection that generated additional genetic gain. The partitioning of genetic trend for trait 1 showed that when we used only the nucleus males in the multiplier (MaleFlow100), the multiplier generated additional gain from two sources. First, compared to the nucleus, the contribution of the nucleus males increased because they contributed through the gene flow and through the selection of multiplier females. Second, the selection of multiplier females contributed as well. When we used both the nucleus males and the multiplier males in the multiplier (MaleFlow20), the multiplier generated further gain through a combination of the sources. First was the contribution of the selection of multiplier females and males. In contrast, the contribution of nucleus selection decreased due to the

323 reduced gene flow. This decrease was due to a smaller number of progeny per nucleus male
324 compared to the MaleFlow100 scenario.

325 On the contrary, trait 2 was not measured in the multiplier and had comparable or smaller
326 genetic trend in the multiplier than in the nucleus. For trait 2 the multiplier animals were
327 selected only on estimated parent average, which resulted in low accuracy selection. In the
328 MaleFlow100 scenario this low accuracy selection resulted in a null contribution of multiplier
329 females to the genetic trend for trait 2 and comparable genetic trends between the nucleus and
330 the multiplier. In the MaleFlow20 scenario with a reduced nucleus-multiplier gene flow this
331 low accuracy selection resulted in the reduced genetic gain for trait 2.

332 **Conclusion**

333 AlphaPart R package is a freely available software for partitioning breeding values and
334 genetic trends. Use of AlphaPart will help breeders to better understand sources of genetic
335 gain and improve their breeding programmes.

336 **Declarations**

337 **Ethics approval and consent to participate**

338 Not applicable

339 **Consent for publication**

340 Not applicable

341 **Availability of data and materials**

342 **Project name:** AlphaPart

343 **Project home page:** <https://cran.r-project.org/package=AlphaPart>

344 **Operating system(s):** Windows, MacOS, Linux

345 **Programming language:** R & C++

346 **License:** GPL-2 | GPL-3

347 **Any restrictions to use by non-academics:** -

348 **Competing interests**

349 Not applicable

350 **Funding**

351 The authors acknowledge support from the BBSRC to The Roslin Institute
352 (BBS/E/D/30002275) and The University of Edinburgh's Data-Driven Innovation
353 Chancellor's fellowship.

354 **Authors' contributions**

355 Not applicable

356 **Acknowledgements**

357 Not applicable

358 **Author's information (optional)**

359

360 **References**

1. Blair HT, Pollak EJ. Estimation of genetic trend in a selected population with and without the use of a control population. *J Anim Sci.* 1984;58:878–86.
2. Sorensen D, Wang C, Jensen J, Gianola D. Bayesian analysis of genetic change due to selection using Gibbs sampling. *Genet Sel Evol.* 1994;26:333.
3. García-Cortés LA, Martínez-Ávila JC, Toro MA. Partition of the genetic trend to validate multiple selection decisions. *Animal.* 2008;2:821–4.
4. Thompson R. Sire Evaluation. *Biometrics.* 1979;35:339–53.
5. Gorjanc G, Potočnik K, García-Cortés LA, Jakobsen J, Dürr J. Partitioning of international genetic trends by origin in Brown Swiss bulls. In: *Proceedings of the 2011 Interbull meeting*; ; August 27-29 2011; Stavanger, Norway. <https://journal.interbull.org/index.php/ib/article/view/1196>.
6. Gorjanc G, Hely F, Amer P. Partitioning international genetic trends by origin in Holstein bulls. Presented at: ICAR 38th Annual Meeting: 28 May - 1 June 2012; Cork, Ireland.
7. Špehar M, Ivkić Z, Bulić V, Barać Z, Gorjanc G. Partitioning of Genetic Trends by Origin in Croatian Simmental Cattle. *Agriculturae Conspectus Scientificus.* 2011;76:301-304.
8. Škorput D, Gorjanc G, Kasap A, Luković Z. Partition of genetic trends by origin in Landrace and Large-White pigs. *animal.* 2015;9:1605–9.
9. Gaynor RC, Gorjanc G, Wilson DL, Money D, Hickey JM. AlphaSimR: An R Package for Breeding Program Simulations. Manuscript in preparation. 2019.
10. Misztal I, Tsuruta S, Strabel T, Auvray B, Druet T, Lee DH. BLUPF90 and related programs (BGF90). In: *Proc 7th World Congress on Genetics Applied to Livestock Production*: August 19-23 2002; Montpellier, France. <http://www.wcgalp.org/system/files/proceedings/2002/blupf90-and-related-programs-bgf90.pdf>.

Figures

Figure 1 Distribution of true breeding values and their partitions by trait, year, and tier in MaleFlow100 scenario.

We show scaled densities of partitions in years 23 and 40 of one simulation replicate. MaleFlow100 uses only nucleus males in the multiplier. Trait 1 is measured in the nucleus and the multiplier, while trait 2 is measured only in the nucleus. Black vertical lines represent the nucleus mean breeding value for a trait in a year.

Figure 2: Partitioning of genetic trend by tier-gender in MaleFlow100 scenario.

The scenario uses nucleus males in the multiplier. Trait 1 is measured in the nucleus and the multiplier, while trait 2 is measured only in the nucleus.

Figure 3 Distribution of true breeding values and their partitions by trait, year, and tier in MaleFlow20 scenario.

We show scaled densities of partitions in years 23 and 40 of one simulation replicate. MaleFlow20 uses nucleus and multiplier males in the multiplier. Trait 1 is measured in the nucleus and the multiplier, while trait 2 is measured only in the nucleus. Black vertical lines represent the nucleus mean breeding value for a trait in a year.

380 **Figure 4 Partitioning of the genetic trend by tier-gender in MaleFlow20 scenario.**

381 The scenarios uses nucleus and multiplier males in the multiplier. Trait 1 is measured in the
382 nucleus and the multiplier, while trait 2 is measured only in the nucleus.