bioRxiv preprint doi: https://doi.org/10.1101/2021.01.21.427651; this version posted January 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

1 **Simulation of sugar kelp (*Saccharina latissima*) breeding guided by practices to**
2 **prioritize accelerated research gains**
3

4 Mao Huang[1]*, Kelly R Robbins[1], Yaoguang Li[2], Schery Umanzor[2,3], Michael Marty-
5 Rivera[2], David Bailey[4], Charles Yarish[2], Scott Lindell[4], Jean-Luc Jannink[1,5]
6

7 [1]Section on Plant Breeding and Genetics, School of Integrative Plant Sciences, Cornell
8 University, Ithaca, NY, USA; [2]Department of Ecology & Evolutionary Biology,
9 University of Connecticut, Stamford CT, USA; [3]College of Fisheries and Ocean
10 Sciences, University of Alaska Fairbanks, Juneau, Alaska; [4]Applied Ocean Physics and
11 Engineering, Woods Hole Oceanographic Institution, Woods Hole, MA, USA; [5]United
12 States Department of Agriculture - Agriculture Research Service, Ithaca, NY, USA
13 *corresponding author: mh865@cornell.edu

bioRxiv preprint doi: https://doi.org/10.1101/2021.01.21.427651; this version posted January 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Abstract

The domestication process of sugar kelp in the Northeast U.S. was initiated by selective breeding two years ago. In this study, we will demonstrate how obstacles for accelerated genetic gain can be assessed using simulation approaches that inform resource allocation decisions in our research. Thus far, we have used 140 wild sporophytes (SPs) that were sampled from the northern Gulf of Maine (GOM) to southern New England (SNE). From these SPs, we sampled gametophytes (GPs) and made and evaluated over 600 progeny SPs from crosses among the GPs. The biphasic life cycle of kelp gives a great advantage in selective breeding as we can potentially select both on the SPs and GPs. However, several obstacles exist, such as the amount of time it takes to complete a breeding cycle, the number of GPs that can be maintained in the lab, and whether positive selection can be conducted on farm tested SPs. Using the GOM population characteristics for heritability and effective population size, we simulated a founder population of 1000 individuals and evaluated the impact of overcoming these obstacles on genetic gain. Our results showed that key factors to improve current genetic gain rely mainly on our ability to induce reproduction of the best farm-tested SPs, and to accelerate the clonal vegetative growth of released GPs so that enough GP biomass is ready for making crosses by the next growing season. Overcoming these challenges could improve rates of genetic gain more than two-fold. Future research should focus on conditions favorable for inducing spring and early summer reproduction, and increasing the amount of GP tissue available in time to make fall crosses.

Keywords

sugar kelp, Saccharina latissima, simulation, breeding, genetic gain, genomic selection

bioRxiv preprint doi: https://doi.org/10.1101/2021.01.21.427651; this version posted January 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

60 **Declarations**
61
62 **Acknowledgements**

63 We acknowledge the funding support from the U.S. Dept. of Energy ARPA-E (DE-
64 AR0000915). We thank Natalie Renier from WHOI Communications Department
65 assisting us to make Figure 1.

66

67 **Competing Interests**

68 All authors of this study declare that there is no conflict of interest in this study.

69

70 **Ethics approval**

71 This research complies with the current laws of the United States of America
72

73 **Consent to participate**

74 Not applicable

75

76 **Consent to publication**

77 All authors read and approved the manuscript for publication

78

79 **Availability of data and material**

80 Supplemental Table is available as Online Resource 1.

81

82 **Code availability**

83 Simulation codes are constructed in R using package AlphaSimR

84

bioRxiv preprint doi: https://doi.org/10.1101/2021.01.21.427651; this version posted January 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

85 **Authors' contributions**

86 MH performed the analyses, wrote the manuscript draft and revised the manuscript
87 together with other coauthors. KRR and J-LJ guided the analyses. J-LJ contributed to
88 analysis scripts. YL, SU, SL and CY helped edit the manuscript. YL, SU, MMR, CY,
89 DB, and SL collected phenotypic data, from which the simulation parameters in this
90 manuscript were estimated. SL, CY, and J-LJ led the project and all authors read and
91 approved the manuscript.
92
93

bioRxiv preprint doi: https://doi.org/10.1101/2021.01.21.427651; this version posted January 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

94 **Introduction**
95
96 Wild kelp forests in the ocean provide important habitat and ecosystem services.
97 They have also been an important source of human food. Due to climate change and other
98 anthropogenic factors, global kelp populations have faced a drastic decline (Moy and
99 Christie 2012, Wernberg et al. 2019, Bricknell et al. 2020). Now kelp farming is largely
100 replacing wild harvests: over 30 million metric tons of seaweed were harvested in 2018,
101 of which 97% came from farms (FAO, 2020). The import of seaweed raw materials to the
102 U.S. in 2016 was more than 10,000 metric tons (over $73 million, National Marine
103 Fisheries Service Office of Science and Technology 2016; Piconi et al. 2020). Uses
104 include human food, animal feed supplements, and pharmaceutical and cosmetic products
105 (Kim et al. 2015; Kim et al. 2017; Kim et al. 2019; Marine Biotech 2015; Schiener et al.
106 2015; Wang et al. 2020; Yarish et al. 2017). Growing kelp biomass in the ocean offers a
107 unique opportunity to avoid many of the challenges associated with terrestrial agriculture
108 systems, particularly the growing competition for arable land and freshwater resources. In
109 order to meet the demand of our growing population by 2050, we must use the oceans
110 responsibly to build a thriving seaweed farming industry for the production of carbon-
111 neutral fuels, biochemicals, animal feed, and food (Capron et al. 2020; Kurt et al. 2020).
112 Kelp cultivation has been established for over 60 years in Asian countries. Most
113 recently, there is growing interest in macroalgal cultivation in Europe, South America,
114 and North America (Buschmann et al. 2017; Grebe et al. 2019; Kim et al. 2019; Geocke
115 et al. 2020). Specifically, there are efforts to selectively breed kelp for large-scale food
116 and bioenergy production (Bjerregaard et al. 2016; Hwang et al. 2019; Valero et al. 2017;
117 Geocke et al. 2020) as well as increased demand for germplasm banking to support future
118 cultivation (Barrento et al. 2016; Wade et al. 2020). The U.S. Department of Energy
119 Advanced Research Projects Agency-Energy (ARPA-E) initiated the Macroalgae
120 Research Inspiring Novel Energy Resources (MARINER) program in order to develop
121 new cultivation, management, and breeding technologies that enable cost-efficient
122 seaweed farming in the large U.S. Exclusive Economic Zone and grow into a global
123 leader in the production of seaweeds. The domestication and breeding of sugar kelp,
124 however, is just beginning.
125 Kelp has a bi-phasic life cycle (Redmond et al. 2014), which provides unique
126 opportunities for selective breeding since breeders could potentially exert selection
127 pressure on both phases within a single growing cycle (Peteiro et al. 2016, Wade et al.
128 2020). Genetic markers have been used in crop breeding for some time, primarily
129 exploiting large marker-trait associations (Bernardo 2016). In the last decade, genomic
130 selection (GS) has been adapted by numerous breeding programs due to its ability in
131 predicting breeding values that are immediately used for making selections (Meuwissen
132 et al. 2001, Jannink et al. 2010). The use of genomic selection in terrestrial agriculture
133 and aquaculture breeding has a track record of improving gains by ~10% per generation
134 (Gjedrem et al. 2012). Genomic selection uses a training population with both phenotypic
135 and genotypic information to build a model, which then can be used to predict the
136 genomic estimated breeding value (GEBV) of individuals that are related to the training
137 population. As the development of genetic markers and genotyping individuals becomes
138 less costly compared to phenotyping, GS allows breeders to make selections more
139 efficiently (Heffner et al. 2010).

bioRxiv preprint doi: https://doi.org/10.1101/2021.01.21.427651; this version posted January 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

140 In 2018, a kelp breeding program was initiated by collecting sporophytes (SPs)
141 from the Gulf of Maine (GOM) to southern New England (SNE). Our primary breeding
142 goal is to improve biomass-related traits including wet weight and percentage dry weight,
143 and to reduce biomass ash content. From the wild-sampled SPs, over 700 uniclonal
144 gametophytes (GPs) were isolated and over 200 of these were grown to sufficient
145 biomass for genotyping and for crossing to create progeny SPs, which were planted and
146 evaluated on nearshore kelp farms. Within a cross, each SP has the exact same genotype,
147 resulting in genetically uniform one-meter line "plots" in the farm. A detailed description
148 is reported in Umanzor et al (2020). In the spring of 2019, the farm-grown SPs were
149 measured, and samples were collected to culture in the lab and induce GPs for the next
150 crossing, planting and harvesting cycle. In our current scheme, we use GS to predict the
151 breeding value of gametophytes, select the best ones, and prioritize crossing these GPs to
152 create new sporophytes. Kelp's biphasic life cycle (Fig. 1a) allows us to potentially exert
153 tremendous selection pressure on GPs, as we aim to predict combining abilities of
154 parental GPs using the SP performance. This will empower us to prioritize crosses and
155 evaluate SPs that are more likely to become high-performing varieties.
156 Given our experience, we now have a better understanding of significant obstacles
157 to our breeding effort and the investments that might be exerted to overcome those
158 obstacles. To guide the research effort objectively, the extent of accelerated gain from
159 different possible investments and interventions needs to be assessed via simulation.
160 These simulations will help early kelp breeding efforts utilize limited research and
161 development investment for maximal breeding efficiency and genetic gain (the
162 improvement in population genetic mean).
163 We have identified four obstacles. First, we can collect sorus tissue in the spring
164 from farm-grown SPs and express meiospores that can be individually isolated to grow
165 out to become clonal GPs. However, thus far, we have not routinely succeeded in
166 producing enough clonal GP biomass to make crosses that can generate hundreds of SPs
167 by the fall of the same year. Instead of completing a breeding cycle in one year, our
168 breeding program started with a two year cycle. This slow growth of the clonal GPs
169 represents ***Obstacle 1***. The technical advancement to overcome *Obstacle 1* and complete
170 a breeding cycle in one year would entail some combination of the following:
171 1. Methods to enhance the growth rate of the GPs so that GPs sampled in the
172 spring have sufficient biomass to make crosses in the fall; or
173 2. Methods to make crosses that require less GP biomass but that
174 nevertheless produce plots with adequate numbers of SPs.
175 Currently, we are limited to making no more than 400 crosses per year, due to the
176 labor intensity of maintaining and growing GP cultures in the lab. This bottleneck limits
177 the number of crosses that can be planted and evaluated and represents ***Obstacle 2***.
178 Limiting the number of crosses and associated phenotypic variance can reduce the
179 expected selection intensity and genetic gain. Overcoming *Obstacle 2* would require the
180 ability to maintain and culture more GPs in the lab.
181 Though we can successfully phenotype and rank SPs after a growing season, we
182 have minimal ability to exert positive selection on them since many of the top ranked SPs
183 did not become reproductive prior to harvest. Consequently, our selection of SPs as
184 parents for the next generation of GPs is limited. The lack of selection pressure
185 applicable to SPs represents ***Obstacle 3***. The biphasic nature of kelp should enable two

bioRxiv preprint doi: https://doi.org/10.1101/2021.01.21.427651; this version posted January 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

186 selection events per breeding cycle, one event on the SPs and one on the GPs they
187 produce. In the absence of selection on the SPs, we currently miss an opportunity for
188 genetic gain. Overcoming *Obstacle 3* would entail rapid identification of top SPs, and
189 artificial laboratory induction of SPs to enter reproductive phase (Pang and Luning,
190 2004).
191 Finally, we have shown that it is possible to automate isolating meiospores
192 individually into 96-well plates using flow cytometry (Augyte et al. 2020). This sorting
193 method showed a maximum effectiveness of 76% in gametophyte development (Augyte
194 et al. 2020). We considered the average value in gametophyte survival (i.e., 24 GPs per
195 plate) as a reference parameter in our breeding program. Low GP survival during flow
196 cytometry represents ***Obstacle 4***. Investment in the flow cytometry method to either
197 increase GP survival or enable the preparation of more plates, thus generating more GPs
198 from which to select, would overcome this fourth obstacle.
199 Using simulation, we aim to compare genetic gain after 5 cycles, examining the
200 impacts of overcoming the aforementioned obstacles. This study will guide our decision-
201 making to optimize resource allocation in the next phase of research, and allow other kelp
202 breeders to focus on advancing these areas most needed.
203 Simulation studies have been a useful tool in assisting breeders' decision-making.
204 They are often used to dissect problems that are difficult (expensive or time consuming)
205 to be addressed experimentally. Simulation models can be used to refine more useful
206 experiments based on prior results and experience. For instance, in order to evaluate
207 different ways of improving nitrogen use efficiency for wheat, Dresbøll and Thorup-
208 Kristensen (2014) simulated models mimicking both above and underground plant and
209 environment interactions as well as effects of crop management strategies. These models
210 provided useful guidelines for crop management and variety selection. Simulation results
211 help optimize breeding resource allocation as researchers compare different strategies and
212 predict the potential effects caused by different variables (Parry et al. 2020, Sun et al.
213 2011, Yamamoto et al. 2016). Simulation approaches were also used to identify the best
214 field experimental design in order to most effectively control for spatial variation in
215 agriculture and forestry studies (Gezan et al. 2010). The selection advantages of GS
216 versus using phenotypic selection were evaluated using simulation approaches for barley
217 (based on real marker data, Iwata and Jannink, 2011) and for *Cryptomeria japonica*
218 (purely simulated data, Iwata et al. 2011). Hickey et al. (2014) simulated breeding
219 schemes incorporating GS and assessed GS accuracies to strategize resources allocated
220 between genotyping versus phenotyping, and between the sizes of populations versus
221 numbers of replications (Lorenz, 2013). The potential genetic gains for a small young
222 sorghum breeding program were assessed via simulation (Muleta et al. 2019).
223 In aquaculture, breeding simulation studies have also been applied to address a
224 variety of questions (Zenger et al. 2019), including assessing the changes of inbreeding
225 rates over time (Bentsen and Olesen 2002), evaluating the effects of mating strategies on
226 the changes of genetic gain in 10 generations of aquaculture selection (Sonesson and
227 Ødegård, 2016), and assessing the genomic prediction accuracy using either identical by
228 state or identical by descent genomic relationship matrices (Vela-Avitúa et al. 2015).
229 Zenger et al. (2019) reported that at least 36 simulation studies were relevant in
230 aquaculture breeding evaluating different mating designs, selection strategy, family and

bioRxiv preprint doi: https://doi.org/10.1101/2021.01.21.427651; this version posted January 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

231 genome sizes, and their effects on changes of breeding program over different
232 generations.
233 For simulation studies to be valuable guides, they must be appropriately
234 parameterized. In our breeding work, we have measured various traits at plot and
235 individual levels (Umanzor et al. 2020). Heritability using data across two growing
236 seasons varied among traits and ranged from 0.05 to 0.58, where dry weight per meter
237 and ash free dry weight heritabilities were approximately 0.4. The percent dry weight had
238 the lowest heritability of 0.05. Furthermore, population genetic analyses on the wild
239 samples were performed to understand their diversity, the relationships among them, and
240 their population history in terms of effective population size (Mao et al. 2020). Our
241 simulation parameters were chosen on the basis of these heritability values and on
242 effective population size estimated using founder markers linkage disequilibrium (LD).
243 In this paper, we present a simulation exercise based on these parameters to prioritize
244 research to overcome the obstacles limiting optimum gain from selection.
245
246 **Materials and Methods**
247 *Defining the four major obstacles*
248 Sampling the kelp sporophytes in the wild and culturing the founder
249 gametophytes (GPs) was a one-time event and is not counted in the breeding cycle. We
250 define a breeding cycle for sugar kelp starting in the fall of the year when we cross GPs,
251 and ending just before we cross GPs for the following year (Fig. 1b).
252 *Obstacle 1* is related to the challenge of cultivating enough biomass from GPs
253 collected from farm-evaluated sporophytes (SPs) to make new crosses within the same
254 breeding cycle. In the simulation, we assumed we could reduce the cycle time from two
255 years to one year. *Obstacle 2* is related to limited capacity to grow GPs for crossing.
256 Simulation scenarios assumed we could design space and labor-saving machines for the
257 lab/hatchery phase and manage higher throughput phenotyping to evaluate 1000 plots
258 instead of 400 plots each year. *Obstacle 3* is based on the fact that we were not able to
259 exert selection on farm grown SPs using their phenotypic data because they were not
260 reproductive and we could not harvest their spores and produce the next generation of
261 GPs. In our simulation, we assume the top-ranked sporophytes could be artificially
262 manipulated to be reproductive, hence we could perform phenotypic selection on these
263 sporophytes rather than applying random selection (Pang and Lüning, 2004). Finally,
264 *Obstacle 4* simply affects how many GPs we can collect per parental SP, with our current
265 maximum of 24 but a possible maximum of 96.
266
267 *Simulation parameters*
268
269 *1. Founder population characteristics*
270 We first needed an estimate of the effective population size of kelp founders. To
271 obtain this estimate, marker data on 140 wild SPs samples from GOM was generated via
272 DArT technology (Mao et al. 2020). Data cleaning was similar to Mao et al. (2020).
273 Markers were filtered by removing ones with more than 10% missing data and those
274 severely departing from Hardy-Weinberg Equilibrium (P-value < 0.01) in more than 25%
275 of the collection sites. Markers with minor allele frequency less than 0.05 and individuals
276 with more than 50% missing data were also removed. A final set of 4906 markers were

bioRxiv preprint doi: https://doi.org/10.1101/2021.01.21.427651; this version posted January 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

277 retained and imputed using the rrBLUP package A.mat function (Endelman et al. 2011)
278 in R (R Development Core Team, 2018). Linkage disequilibrium between markers was
279 estimated using the genetics package (Warnes et al, 2012). Average LD score were
280 estimated to be 0.08, which were then used in estimating the effective sample size (N_e),
281 according to Sved (1971):
282

$$E(r^2) \approx \frac{1}{1 + 4N_e c}$$

283
284 where $E(r^2)$ is the expected r^2 for which we used the average LD score of 0.08, and c is
285 the recombination rate among all sites assumed to be 0.5, given that the vast majority of
286 pairs of sites are on different chromosomes. This gives an estimated $N_e = 60$. We know
287 that the GOM population is strongly structured (Mao et al. 2020), which may cause N_e to
288 be underestimated. Thus we also ran simulations with a setting of $N_e = 600$. A total of
289 1000 SP individuals were simulated as our founders with the effective population size of
290 either $N_e = 60$ or $N_e = 600$.
291 The ploidy level was set to 2 and the number of chromosomes was assumed to be
292 31 based on its close congener *Saccharina japonica* (Liu et al. 2012). Per chromosome,
293 the number of segregating sites and the number of QTL were set to 500 and 100,
294 respectively. These values assume that the trait is polygenic but are otherwise somewhat
295 arbitrary and chosen referring to those in Muleta et al. (2019).
296 A mixed model, including genetic effects of SPs as random effects, then growth
297 line, blocks, date of harvesting, and reference checks as the fixed effects, was conducted
298 to estimate the narrow sense heritability using 2018-2019 and 2019-2020 two field season
299 GOM farm SP data. Heritability was estimated using:

300 $$h^2 = \frac{\sigma_A^2}{\sigma_A^2 + \sigma_E^2}$$

301 where σ_A^2 is the estimated additive variance for the SPs and σ_E^2 is error variance from a
302 mixed model. Trait heritabilities ranged from 0.05 to 0.50 for plot-level traits and 0.06 to
303 0.58 for individual-level traits using both years' data. In the simulation study, the trait
304 genotypic variance was set at 1 and error variance at 4 or 1 so that initial heritability was
305 0.20 or 0.50 for biomass related traits.
306
307 *2. Breeding Scheme*
308 Historically, we have been able to produce enough GP biomass to make 2 crosses
309 per GP. Consequently, we assumed that same capacity in the simulation scheme. We
310 created initial SP founder populations of 1000 individuals and allowed each SP to
311 generate two GPs, giving enough GPs to make either 400 or 1000 crosses for downstream
312 generations without exerting selection pressure on the founder population. The simulation
313 program randomly assigned "F" and "M" sexes to GPs generated from the founder
314 population. Ten percent of SPs were selected either randomly or based on phenotype to
315 be parents of the next generation GPs. Thus, 40 and 100 SPs were selected from 400 and
316 1000 SPs evaluated, respectively. From these selected SPs, we assumed flow cytometry
317 would be used to obtain GPs from each SP (Augyte et al. 2020). This automated spore
318 sorting technology produces viable uni-clonal isolations on average in 25% of the wells
319 of a 96-well plate (i.e., 24 GPs) from spores released from an individual SP. An ideal

bioRxiv preprint doi: https://doi.org/10.1101/2021.01.21.427651; this version posted January 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

320 situation where all 96 GPs in the plate are viable was also included in the simulation
321 scheme. Generating 96 GPs means there will be four times more GPs to select from to
322 make crosses for SPs, enabling higher selection intensity. The farm-evaluated SPs
323 phenotypic data from both years was used to train a genomic selection model which was
324 used to predict the breeding values of all GPs coming out of the flow cytometry process.
325 Either 200 or 500 top ranked GPs based on their predicted breeding values would be
326 selected to make the 400 or 1000 crosses as farm-evaluated SPs plots. Note that with this
327 scheme, changing the number of SPs evaluated does not change the selection intensity
328 either during SP or GP selection, whereas changing the number of GPs generated changes
329 the selection intensity during the GP selection stage.
330
331 *3. Estimating Genetic gain over 5 breeding cycles*
332 Breeding scheme simulation was done using the AlphaSimR package in R (Faux
333 et al. 2016). Each scheme was simulated 100 times, and the average genetic gain as well
334 as genetic variance at each GP stage was calculated over 5 cycles of selection. Because
335 we were mainly interested in evaluating the trend of genetic gain from different breeding
336 schemes, the reference point for genetic gain could either be for GPs or SPs. We used
337 GPs.
338
339 **Results**
340
341 *Simulation output*
342 The ability to exert selection on the farm-evaluated SPs (SelectSP), the number of
343 years per breeding cycle time (CycleTime), and the number of gametophytes per SP
344 surviving the flow cell cytometry system (nGP) were the three significant contributors to
345 the changes of genetic mean over time (Table1). We did not observe significant
346 interactions between these factors (Table 1).
347 The baseline simulation scheme represented our current state of the art, where 400
348 SPs are phenotyped in the field, no selection pressure is exerted on the SPs, and the
349 breeding cycle takes 2 years. Our number of GPs per SP was either 24 or 96. The changes
350 from Fig. 2a to Fig. 2b reflect the effects of overcoming *Obstacle 4* where higher nGP
351 could be obtained through a single cell sorting flow cytometry step (nGP=24 versus
352 nGP=96). This change from nGP=24 to nGP=96 led to a gain increase of 37% averaged
353 across all other factors (nGP, Table 1, Fig. 2). Relative to the baseline, the ability to exert
354 selection on SPs (*Obstacle 3*) and decreasing the breeding cycle time (*Obstacle 1*) led to
355 gain increases of 101% and 45%, respectively, averaged across all other factors. Though
356 the effect of increasing the number of plots phenotyped was not statistically significant
357 (NumPlots, Table 1), numerically this change increased gain by an average of 11%
358 (overcoming *Obstacle 2*). We did not observe significant interactions: the effects of
359 overcoming each obstacle were additive (Table 1), and overcoming all four obstacles led
360 to the greatest gain (Fig. 2). Heritability also played a role in affecting the genetic gain
361 (Table 1), where h^2=0.5 generated higher genetic mean after 7 years of breeding than
362 h^2=0.2 (Figs. 2a and 2b). This trend was consistent regardless of the number of
363 gametophytes or effective population size.
364 The breeding scheme interventions simulated also affected the genetic variance
365 remaining after seven years of improvement (Fig. 3). All three interventions that

bioRxiv preprint doi: https://doi.org/10.1101/2021.01.21.427651; this version posted January 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

366 significantly increased genetic gain also caused decreases in genetic variance. The
367 smallest change in genetic variance occurred as a result of selecting SPs on phenotype.
368 Selection causes variance decreases both because of the Bulmer effect and because high
369 fitness ancestors contribute disproportionately to descendants. With the 1-year per cycle
370 scheme, the population went through twice as many selection events as with the 2-year
371 scheme, leading to a greater decrease in genetic variation over the seven years (Fig. 3,
372 Online Resource 1). For all combinations of other factors, there was a higher final genetic
373 variance when nGP was 24 than when it was 96. The increased selection intensity from
374 this intervention caused a greater variance decrease than for any other intervention. The
375 only intervention that caused increased final genetic variance was evaluating more SP
376 plots per year (1000 versus 400). In this case, increasing the number of plots caused
377 increased effective population size and thus greater maintenance of variance. It also
378 caused increased genomic prediction accuracy, which has also been shown to maintain
379 genetic variance (Jannink et al. 2010). These effects are depicted in Figure 4. The low
380 level of interaction between simulated factors can also be seen in Figure 4 by the fact that
381 lines linking simulation settings with and without the interventions are approximately
382 parallel and of similar length, indicating that changing one factor has basically the same
383 effect regardless of the levels of the other factors.
384
385
386 **Discussion**
387
388 *Which obstacle should we focus on*?
389 Simulation is a useful tool for guiding researchers in decision-making especially
390 for young breeding programs (Muleta et al. 2019) and in assisting selection strategy and
391 decision-making processes (Zenger et al. 2019). We simulated different breeding
392 schemes, each overcoming a major obstacle we have encountered in two seasons of kelp
393 breeding. Assessing which scheme generated the highest genetic gain allows us to
394 prioritize research directions and derive the most benefit from a limited budget.
395 The simulation revealed a robust result that the highest genetic gain can be
396 achieved by exerting selection on SPs phenotypically (overcoming *Obstacle 3*), and then
397 by reducing the time for obtaining sufficient GP biomass such that a one-year cycle is
398 enabled compared to our current two-year cycle (overcoming *Obstacle 1*). Increasing the
399 number of viable GP we obtain per parental SP (overcoming *Obstacle 4*) also delivered
400 significant gain, while, somewhat surprisingly, phenotyping more SP plots (overcoming
401 *Obstacle 2*) did not. These conclusions were not affected by the founder population
402 effective population size or the trait heritability. Thus, the clear direction to prioritize
403 breeding enhancement is to induce SP spore release and to modify GP culture to
404 accelerate growth. In addition, we should experiment with the amount of GP biomass
405 needed to make sufficient SP progeny. We may not need a full one-meter of line we use
406 as an evaluation plot.
407 Developing the ability to induce top performing SPs to release spores (*Obstacle 3*)
408 can be a challenge for the following reasons. First, typically only approximately 10% of
409 the plots are fertile at the optimum time of harvest as measured by most marketable yield.
410 These crosses are not necessarily the top performing ones. Ideally, we aim to select
411 crosses in the top 10% for performance and artificially induce them indoors if necessary.

bioRxiv preprint doi: https://doi.org/10.1101/2021.01.21.427651; this version posted January 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

412 This has proven successful on a small scale if desirable SPs are identified within a day of
413 harvest. Overcoming *Obstacle 3* requires greater investment in labor to identify and
414 separate candidate SPs, and investment in culture space.
415 Our second best option is to accelerate GP growth by overcoming *Obstacle 1*,
416 which could also be the hardest task. In brief, it takes four to eight weeks to induce
417 immature SPs to full maturity and release meiospores under artificial conditions in the lab
418 (Pang and Lining 2004, Flavin et al. 2013, Remond et al. 2014). Once meiospores are
419 released, flow-cytometry techniques can be implemented to isolate single-cell
420 gametophytes into 96-well plates. A second isolation is performed approximately two to
421 four months later when GPs develop into tufts large enough (>100 μm), to be sexed and
422 moved to individual Petri dishes for filament fragmentation. Once sufficient uniclonal
423 biomass is achieved (~10 mg to cover 1 m plots), which can take up to another four
424 months, crosses are made by mixing female and male GPs at a 2:1 ratio (Umanzor et al.
425 2020, Fig. 1). Outplanting at sea occurs 4-6 weeks following SP attachment onto the seed
426 string (Flavin et al. 2013, Remond et al. 2014). Overall, this process of uniclonal GP
427 isolation, growth and crossing is effective but typically requires 12 months, in contrast to
428 the six months between optimal kelp harvesting (end of May to early June) to crossing
429 and outplanting (November to December).
430 Possible means of accelerating GP growth include optimizing lighting, nutrient
431 and temperature regimes, as well as novel biomass fragmentation protocols. It might be
432 possible to optimize GP biomass development by transferring them earlier to plates with
433 bigger wells (i.e. from 96-well plates to 24-well plates) that would allow better light
434 penetration. Generally, GP growth is limited by the natural biological programming of
435 cell division and a propensity to self-shade in its puff-ball growth form. However, some
436 GPs grow faster than others, and selecting for GP growth performance could be
437 incorporated in the breeding program. In order to test and see if 1-year cycle time is
438 feasible in our current breeding program (approximately six months between GP isolation
439 and crossing), we are experimenting with using a minimum amount of biomass to make
440 crosses and generate at least a single SP blade. The function of this blade would not be
441 for evaluation of SP performance but for recombining the best GPs in the hope of getting
442 improved recombinants. The approach will generate phenotypic data on the individual SP
443 but not on biomass per meter of line, which is a plot-level trait. Hence this procedure
444 would not be a full representation of the simulated 1-year per cycle scheme. Nonetheless,
445 this will be a proof of concept for us to accelerate the GP culturing process.
446 Another possibility that is used in forage breeding (Resende et al. 2013) would be
447 to evaluate segregating plots, in our case created from crossing multiple female GPs from
448 one SP with multiple male GPs from another SP. The between plots variance for such
449 mixed plots would be less than that for the uniform SPs plots. Furthermore, maintaining
450 multiple individual GPs only until they can be sexed and co-cultured together would
451 reduce some labor. Such mixed plots would generate sufficient biomass more quickly to
452 facilitate one-year breeding cycles.
453 Overcoming *Obstacle 2* by increasing the number of GPs per parental SP can
454 potentially be done easily. A simple approach would be to increase the number of plates
455 automatically sorted by flow cytometry per parental SP, which would increase the
456 number of GPs in the genomic selection step, allowing higher selection intensity.
457 Nonetheless, this would result in increasing the number of cultures to maintain in the lab,

bioRxiv preprint doi: https://doi.org/10.1101/2021.01.21.427651; this version posted January 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

458 which leads to more labor and cost. The use of flow cytometry sorting expedites the
459 initial isolation process but the parameters determine the survival of spores is not well
460 understood. The condition of sorus tissue prior to spore release and sorting likely has an
461 effect on spore viability. Percentage viability varied across samples presumably because
462 of differences in sorus tissue condition and handling prior to sorting (Augyte et al, 2020).
463 An issue that should be investigated is whether the selection pressure caused by flow
464 cytometry mortality has pleiotropic effects that might negatively affect SP growth or
465 reproduction. If not, the mortality should generate its own natural selection response that
466 will eventually mitigate this obstacle.
467 Increasing the number of plots (from 400 to 1000) could be accomplished without
468 new research, but could be costly since it would require more GP grow out space and
469 labor. This change generated only a small increase in the rate of genetic gain. An
470 additional benefit to increasing the number of SPs being phenotyped, however, was that it
471 maintained genetic diversity and slowed down the decrease of genetic variance (Fig. 3,
472 Fig. 4). The proportion of GPs selected out of SPs were the same regardless of testing
473 400 or 1000 plots, hence increasing the number of plots did not change the selection
474 differential. It did, however, affect the training population size of GS models when
475 selecting new generations of GPs. Larger training population size usually contributes to
476 increased GS accuracy (Poland et al. 2012; Huang et al. 2016). In this case, the increased
477 phenotypic data led to an improved genomic prediction model and its ability to
478 distinguish among-family versus within family effects. That ability can decrease the co-
479 selection of relatives leading to greater maintenance of genetic variation (Jannink et al.
480 2010). Interestingly, every intervention that led to greater genetic gain also led to greater
481 loss of genetic variance for all changes in practice (Selection on SP, Cycle Time, nGP per
482 parental SP), *except* increasing the number of phenotyped plots which had both increased
483 gain and decreased variance lost (Fig. 4). We also observed in some cases that the
484 principal effect of increasing the number of plots was to cause greater variance to be
485 retained, without increasing the gain from selection substantially (in Fig. 4 the gray lines
486 were close to vertical). Hence, it seems likely that this intervention would benefit our
487 breeding program over the long term.
488 While in this discussion we have treated heritability as fixed, that is not strictly
489 true. Heritability might be increased if we could improve our planting technique to ensure
490 that plots are more uniformly covered by SPs, so that we obtain successful and uniform
491 growth of SPs in the field. Not surprisingly, higher heritability leads to greater final gain
492 (Figs. 2 and 3). The decreasing trend of genetic variance was expected leading to a
493 relationship where higher final genetic gain coincided with lower genetic variance. It is
494 important to maintain the diversity while we improve the progeny performance (Heffner
495 et al. 2009; Lin et al. 2016). Overall, the robustness of these simulation findings should
496 give us confidence in the research directions they suggest. We believe that these priorities
497 will greatly help accelerate genetic gain in breeding programs and therefore increase the
498 value of kelp farming in the United States and globally.

bioRxiv preprint doi: https://doi.org/10.1101/2021.01.21.427651; this version posted January 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

499
500 **Reference List**
501

502 Augyte S, Wikfors GH, Pitchford S, Marty-Rivera M, Umanzor S, Lindell S,
503 Bailey D, Yarish C (2020) The application of flow cytometry for kelp meiospore
504 isolation. Algal Research 46:101810; https://doi.org/10.1016/j.algal.2020.101810
505 Barrento S, Camus C, Sousa-Pinto I, Buschmann AH (2016) Germplasm banking
506 of the giant kelp: Our biological insurance in a changing environment. Algal Research
507 13:134-140
508 Bentsen HB, Olesen I (2002) Designing aquaculture mass selection programs to
509 avoid high inbreeding rates. Aquaculture 204:349-359
510 Bernardo R (2016) Bandwagons I, too, have known. Theor Appl Genet 129:2323-
511 2332
512 Bjerregaard R, Valderrama D, Radulovich R, Diana J, Capron M, Mckinnie CA,
513 Cedric M, Hopkins K, Yarish C, Goudey C (2016) Seaweed aquaculture for food
514 security, income generation and environmental health in tropical developing countries.
515 World Bank Group, Washington, D.C. URL
516 http://documents.worldbank.org/curated/en/947831469090666344/Seaweed-aquaculture-
517 for-food-security-income-generation-and-environmental-health-in-Tropical-Developing-
518 Countries;jsessionid=4sLY8b149Hwa-8ramT5do35G (Report #107147).
519 Bricknell IR, Birkel SD, Brawley SH, Van Kirk T, Hamlin HJ, Capistrant-Fossa
520 K, Huguenard K, Van Walsum GP, Liu ZL, Zhu LH (2020) Resilience of cold water
521 aquaculture: a review of likely scenarios as climate changes in the Gulf of Maine.
522 Reviews in Aquaculture 13:460-503
523 Broch OJ, Slagstad D (2012) Modelling seasonal growth and composition of the
524 kelp Saccharina latissima. J Appl Phycol 24:759-776
525 Buschmann AH, Camus C, Infante J, Neori A, Israel Á, Hernández-González
526 MC, Pereda SV, Gomez-Pinchetti JL, Golberg A, Tadmor-Shalev N (2017) Seaweed
527 production: overview of the global state of exploitation, farming and emerging research
528 activity. Eur J Phycol 52:391-406
529 Capron ME, Stewart JR, de Ramon N'Yeurt A, Chambers MD, Kim JK, Yarish
530 C, Jones AT, Blaylock RB, James SC, Fuhrman R (2020) Restoring pre-industrial CO2
531 levels while achieving Sustainable Development Goals. Energies 13:4972
532 Dresbøll DB, Thorup-Kristensen K (2014) Will breeding for nitrogen use efficient
533 crops lead to nitrogen use efficient cropping systems?: a simulation study of G× E× M
534 interactions. Euphytica 199:97-117
535 Endelman JB (2011) Ridge regression and other kernels for genomic selection
536 with R package rrBLUP. The Plant Genome 4:250-255
537 FAO (2020) The state of worldwide fisheries and aquaculture. URL
538 http://www.fao.org/3/ca9229en/ca9229en.pdf (Accessed Dec 22, 2020)
539 Faux AM, Gorjanc G, Gaynor RC, Battagin M, Edwards SM, Wilson DL, Hearne
540 SJ, Gonen S, Hickey JM (2016) AlphaSim: software for breeding program simulation.
541 The plant genome 9:1-14
542 Flavin K, Flavin N, Flahive B (2013) Kelp farming manual: a guide to the
543 processes, techniques, and equipment for farming kelp in New England waters. Ocean
544 Approved LLC, Portland. URL

bioRxiv preprint doi: https://doi.org/10.1101/2021.01.21.427651; this version posted January 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

545 https://static.squarespace.com/static/52f23e95e4b0a96c7b53ad7c/t/52f78b0de4b0374e6a
546 0a4da8/1391954701750/ OceanApproved_KelpManualLowRez.pdf (Accessed Dec 22,
547 2020)
548 Gezan SA, White TL, Huber DA (2010) Accounting for spatial variability in
549 breeding trials: A simulation study. Agron J 102:1562-1571
550 Gjedrem T, Robinson N, Rye M (2012) The importance of selective breeding in
551 aquaculture to meet future demands for animal protein: a review. Aquaculture 350:117-
552 129
553 Goecke F, Klemetsdal G, Ergon Å (2020) Cultivar Development of Kelps for
554 Commercial Cultivation—Past Lessons and Future Prospects. Frontiers in Marine
555 Science 7:110
556 Grebe GS, Byron CJ, Gelais AS, Kotowicz DM, Olson TK (2019) An ecosystem
557 approach to kelp aquaculture in the Americas and Europe. Aquaculture Reports
558 15:100215
559 Heffner EL, Lorenz AJ, Jannink JL, Sorrells ME (2010) Plant breeding with
560 genomic selection: gain per unit time and cost. Crop Sci 50:1681-1690
561 Heffner EL, Sorrells ME, Jannink JL (2009) Genomic selection for crop
562 improvement. Crop Sci 49:1-12
563 Hickey JM, Dreisigacker S, Crossa J, Hearne S, Babu R, Prasanna BM, Grondona
564 M, Zambelli A, Windhausen VS, Mathews K (2014) Evaluation of genomic selection
565 training population designs and genotyping strategies in plant breeding programs using
566 simulation. Crop Sci 54:1476-1488
567 Huang M, Cabrera A, Hoffstetter A, Griffey C, Van Sanford D, Costa J,
568 McKendry A, Chao S, Sneller C (2016) Genomic selection for wheat traits and trait
569 stability. Theor Appl Genet 129:1697-1710
570 Hwang EK, Yotsukura N, Pang SJ, Su L, Shan TF (2019) Seaweed breeding
571 programs and progress in eastern Asian countries. Phycologia 58:484-495
572 Iwata H, Hayashi T, Tsumura Y (2011) Prospects for genomic selection in conifer
573 breeding: a simulation study of Cryptomeria japonica. Tree Genet Genom 7:747-758
574 Iwata H, Jannink JL (2011) Accuracy of genomic selection prediction in barley
575 breeding programs: a simulation study based on the real single nucleotide polymorphism
576 data of barley breeding lines. Crop Sci 51:1915-1927
577 Jannink J-L, Lorenz AJ, Iwata H (2010) Genomic selection in plant breeding:
578 from theory to practice. Briefings in functional genomics 9:166-177
579 Kim J, Stekoll M, Yarish C (2019) Opportunities, challenges and future directions
580 of open-water seaweed aquaculture in the United States. Phycologia 58:446-461
581 Kim JK, Kraemer GP, Yarish C (2015) Use of sugar kelp aquaculture in Long
582 Island Sound and the Bronx River Estuary for nutrient extraction. Mar Ecol Prog Ser
583 531:155-166
584 Kim JK, Yarish C, Hwang EK, Park M, Kim Y, Kim JK, Yarish C, Hwang EK,
585 Park M, Kim Y (2017) Seaweed aquaculture: cultivation technologies, challenges and its
586 ecosystem services. Algae 32:1-13
587 Kurt T, Vijn S, Compart DP, Dutta N, Foukis A, Hess M, Hristov AN, Kalscheur
588 K, Kebreab E, Nuzhdin S (2020) Key Considerations for the Use of Seaweed to Reduce
589 Enteric Methane Emissions from Cattle. Frontiers in Veterinary Science 7:1135

bioRxiv preprint doi: https://doi.org/10.1101/2021.01.21.427651; this version posted January 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

590 Lin Z, Cogan NO, Pembleton LW, Spangenberg GC, Forster JW, Hayes BJ,
591 Daetwyler HD (2016) Genetic gain and inbreeding from genomic selection in a simulated
592 commercial breeding program for perennial ryegrass. The Plant Genome 9:1-12
593 Liu Y, Bi Y, Gu J, Li L, Zhou Z (2012) Localization of a female-specific marker
594 on the chromosomes of the brown seaweed Saccharina japonica using fluorescence in situ
595 hybridization. PLoS One 7:e48784
596 Lorenz AJ (2013) Resource allocation for maximizing prediction accuracy and
597 genetic gain of genomic selection in plant breeding: a simulation experiment. G3: Genes,
598 Genomes, Genetics 3:481-491
599 Marine Biotech (2015) Examples of Marine Biotechnology successes. URL
600 www.marinebiotech.eu/wiki/Examples_of_Marine_Biotechnology_successes (Accessed
601 Dec 22, 2020)
602 Mao X, Augyte S, Huang M, Hare MP, Bailey D, Umanzor S, Marty-Rivera M,
603 Robbins KR, Yarish C, Lindell S (2020) Population Genetics of Sugar Kelp Throughout
604 the Northeastern United States Using Genome-Wide Markers. Frontiers in Marine
605 Science 7:694
606 Meuwissen T, Hayes B, Goddard M (2001) Prediction of total genetic value using
607 genome-wide dense marker maps. Genetics 157:1819-1829
608 Moy FE, Christie H (2012) Large-scale shift from sugar kelp (Saccharina
609 latissima) to ephemeral algae along the south and west coast of Norway. Mar Biol Res
610 8:309-321
611 Muleta KT, Pressoir G, Morris GP (2019) Optimizing genomic selection for a
612 sorghum breeding program in Haiti: a simulation study. G3: Genes, Genomes, Genetics
613 9:391-401
614 National Marine Fisheries Service Office of Science and Technology (2016) URL
615 https://www.fisheries.noaa.gov/about/office-science-and-technology (Accessed Dec 22,
616 2020)
617
618 Pang SJ, Lüning K (2004) Breaking seasonal limitation: year-round sporogenesis
619 in the brown alga Laminaria saccharina by blocking the transport of putative sporulation
620 inhibitors. Aquaculture 240:531-541
621 Parry H, Kalyebi A, Bianchi F, Sseruwagi P, Colvin J, Schellhorn N, Macfadyen
622 S (2020) Evaluation of cultural control and resistance-breeding strategies for suppression
623 of whitefly infestation of cassava at the landscape scale: a simulation modeling approach.
624 Pest Manage Sci 76(8):2699–2710
625 Peteiro C, Sánchez N, Martínez B (2016) Mariculture of the Asian kelp Undaria
626 pinnatifida and the native kelp Saccharina latissima along the Atlantic coast of Southern
627 Europe: An overview. Algal research 15:9-23
628 Piconi, P., Veidenheimer, R., & Chase B. (2020) Edible Seaweed Market
629 Analysis. Island Institute, Maine, US. URL : http://www.islandinstitute.org/edible-
630 seaweedmarket-analysis-2020 (Accessed Dec 22, 2020)
631 Poland J, Endelman J, Dawson J, Rutkoski J, Wu S, Manes Y, Dreisigacker S,
632 Crossa J, Sánchez-Villeda H, Sorrells M (2012) Genomic selection in wheat breeding
633 using genotyping-by-sequencing. The Plant Genome 5:103-113
634

bioRxiv preprint doi: https://doi.org/10.1101/2021.01.21.427651; this version posted January 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

635 R Development Core Team (2018) R: A language and environment for statistical
636 computing. The R Foundation for Statistical Computing, Vienna, Austria. URL
637 http://www.R-project.org/
638 Redmond, S., Green, L., Yarish, C., Kim, J., & Neefus, C. (2014) New England
639 seaweed culture handbook-nursery systems. Connecticut Sea Grant, 93 pp. URL
640 https://seagrant.uconn.edu/2014/01/01/new-england-seaweed-culture-handbook-nursery-
641 systems/ (Accessed Dec 22, 2020)
642 Resende RMS, Casler MD, de Resende MDV (2013) Selection methods in forage
643 breeding: a quantitative appraisal. Crop Sci 53:1925-1936
644 Schiener P, Black KD, Stanley MS, Green DH (2015) The seasonal variation in
645 the chemical composition of the kelp species Laminaria digitata, Laminaria hyperborea,
646 Saccharina latissima and Alaria esculenta. J Appl Phycol 27:363-373
647 Sonesson AK, Ødegård J (2016) Mating structures for genomic selection breeding
648 programs in aquaculture. Genet Sel Evol 48:46
649 Sun X, Peng T, Mumm R (2011) The role and basics of computer simulation in
650 support of critical decisions in plant breeding. Mol Breed 28:421-436
651 Sved J (1971) Linkage disequilibrium and homozygosity of chromosome
652 segments in finite populations. Theor Popul Biol 2:125-141
653 Umanzor, S., Li, Y., Bailey, D., Augyte, S., Marty-Rivera, M., Huang, M.,
654 Jannink, J-L, Yarish, C., Lindell, S. (2020) Comparative Analysis of Morphometric Traits
655 of Strains of Sugar and Skinny Kelp, Saccharina spp., from the Northwest Atlantic Coast.
656 Submitted Journal of World Aquaculture Society. December 2020
657 Valero M, Guillemin M-L, Destombe C, Jacquemin B, Gachon CM, Badis Y,
658 Buschmann AH, Camus C, Faugeron S (2017) Perspectives on domestication research for
659 sustainable seaweed aquaculture. Perspectives in Phycology 4:33-46
660 Vela-Avitúa S, Meuwissen TH, Luan T, Ødegård J (2015) Accuracy of genomic
661 selection for a sib-evaluated trait using identity-by-state and identity-by-descent
662 relationships. Genet Sel Evol 47:9
663 Wade R, Augyte S, Harden M, Nuzhdin S, Yarish C, Alberto F (2020) Macroalgal
664 germplasm banking for conservation, food security, and industry. PLoS Biol
665 18:e3000641
666 Wang X, Yao J, Zhang J, Duan D (2020) Status of genetic studies and breeding of
667 Saccharina japonica in China. Journal of Oceanology and Limnology 38:1064-1079
668 Warnes G, Leisch F, Man M, Warnes MG (2012) Package 'genetics'. Rochester,
669 NY
670 Wernberg T, Krumhansl K, Filbee-Dexter K, Pedersen MF (2019) Status and
671 trends for the world's kelp forests. World seas: An environmental evaluation. Elsevier,
672 pp 57-78
673 Yamamoto E, Matsunaga H, Onogi A, Kajiya-Kanegae H, Minamikawa M,
674 Suzuki A, Shirasawa K, Hirakawa H, Nunome T, Yamaguchi H (2016) A simulation-
675 based breeding design that uses whole-genome prediction in tomato. Scientific reports
676 6:19454
677 Yarish C, Kim JK, Lindell S, Kite-Powell H (2017) Developing an
678 environmentally and economically sustainable sugar kelp aquaculture industry in
679 southern New England: from seed to market. *EEB Articles* 38. URL
680 http://opencommons.uconn.edu/marine_sci/4 (Accessed Dec 22, 2020)

bioRxiv preprint doi: https://doi.org/10.1101/2021.01.21.427651; this version posted January 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

681 Zenger KR, Khatkar MS, Jones DB, Khalilisamani N, Jerry DR, Raadsma HW
682 (2019) Genomic selection in aquaculture: application, limitations and opportunities with
683 special reference to marine shrimp and pearl oysters. Frontiers in genetics 9:693
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706

bioRxiv preprint doi: https://doi.org/10.1101/2021.01.21.427651; this version posted January 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Tables and Figure legends

Table 1. ANOVA on genetic mean split by founder effective population size (N_e) and heritability (h^2).

a. N_e = 60, h^2 = 0.5

	Df	Sum Sq	Mean Sq	*F*	*P-value*
SelectSP†	1	47.6	47.6	22.4	7.34E-06***
NumPlots	1	1.0	1.0	0.5	0.496
CycleTime	1	13.7	13.7	6.4	0.013*
nGP	1	10.0	10.0	4.7	0.032*
SelectSP:NumPlots	1	0.0	0.0	0.0	0.927
SelectSP:CycleTime	1	0.5	0.5	0.3	0.613
SelectSP:nGP	1	0.4	0.4	0.2	0.648
NumPlots:CycleTime	1	0.1	0.1	0.0	0.876
NumPlots:nGP	1	0.1	0.1	0.0	0.830
CycleTime:nGP	1	0.4	0.4	0.2	0.663
Residuals	101	215.0	2.1		

b. N_e = 600, h^2 = 0.5

	Df	Sum Sq	Mean Sq	*F*	*P-value*
SelectSP†	1	48.5	48.5	31.1	2.02E-07***
NumPlots	1	0.4	0.4	0.3	0.614
CycleTime	1	10.8	10.8	6.9	0.010**
nGP	1	6.9	6.9	4.4	0.038*
SelectSP:NumPlots	1	0.0	0.0	0.0	0.897
SelectSP:CycleTime	1	0.4	0.4	0.3	0.610
SelectSP:nGP	1	0.4	0.4	0.3	0.614
NumPlots:CycleTime	1	0.0	0.0	0.0	0.902
NumPlots:nGP	1	0.1	0.1	0.0	0.838
CycleTime:nGP	1	0.3	0.3	0.2	0.666
Residuals	101	157.2	1.6		

c. N_e = 60, h^2 = 0.2

	Df	Sum Sq	Mean Sq	*F*	*P-value*
SelectSP†	1	19.3	19.3	15.4	0.000***
NumPlots	1	1.2	1.2	0.9	0.339
CycleTime	1	8.1	8.1	6.5	0.012*
nGP	1	6.7	6.7	5.3	0.023*
SelectSP:NumPlots	1	0.0	0.0	0.0	0.967
SelectSP:CycleTime	1	0.2	0.2	0.1	0.726
SelectSP:nGP	1	0.2	0.2	0.2	0.660
NumPlots:CycleTime	1	0.0	0.0	0.0	0.849

bioRxiv preprint doi: https://doi.org/10.1101/2021.01.21.427651; this version posted January 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

NumPlots:nGP	1	0.1	0.1	0.1	0.733
CycleTime:nGP	1	0.2	0.2	0.2	0.676
Residuals	101	126.6	1.3		

d. $N_e = 600$, $h^2 = 0.2$

	Df	Sum Sq	Mean Sq	*F*	*P-value*
SelectSP†	1	19.4	19.4	21.4	1.11E-05***
NumPlots	1	0.4	0.4	0.5	0.486
CycleTime	1	6.8	6.8	7.6	0.007**
nGP	1	4.3	4.3	4.7	0.032*
SelectSP:NumPlots	1	0.0	0.0	0.0	0.897
SelectSP:CycleTime	1	0.3	0.3	0.3	0.599
SelectSP:nGP	1	0.1	0.1	0.2	0.691
NumPlots:CycleTime	1	0.0	0.0	0.0	0.827
NumPlots:nGP	1	0.1	0.1	0.1	0.769
CycleTime:nGP	1	0.1	0.1	0.1	0.726
Residuals	101	91.4	0.9		

* P<0.05,** P<0.001, *** P<0.0001

† SelectSP: Selection among SP based on phenotype or at random. NumPlots: Common garden of 400 versus 1000 field plots. CycleTime: 1-year versus 2-year cycle. nGP: number of GPs obtained per parental SP of 24 or 96.

bioRxiv preprint doi: https://doi.org/10.1101/2021.01.21.427651; this version posted January 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Figure 1.
(a.) Biphasic life cycle and breeding pipeline of sugar kelp (*S. latissima)* in our research project. Represented are meiospore release, flow cell sorting to 96-well plates, propagation to sufficient biomass for crossing, spraying of crossed SPs onto seed string, and outplanting to a farm-like common garden field experiment.
(b.) Breeding scheme timeline view and the corresponding obstacles on number of GPs (nGP), Number of SP plots evaluated on farm (NumPlots), the selection on SPs (phenotypic vs random selection), and CycleTime (1-year vs 2-year).

Figure 2.
Genetic mean from different breeding schemes over 7 years. The routine breeding scheme starts in year 3. Each figure shows NumPlots: Evaluate 400 versus 1000 plots; SelectSP: phenotypically select the best (pheno) versus random (rand) sporophytes for producing new crosses; and CycleTime: 1-year (1yr) versus 2-year (2yr). Subpanels separate different founder population effective population sizes of 60 (Ne60) and 600 (Ne600) and trait heritabilities of h^2 =0.5 and h^2 =0.2 when a.) 24 or b.) 96 gametophytes were propagated from each parental SP. Each scheme was repeated 100 times and genetic values shown were averages. The standard error was smaller than the figure symbols and is not shown.

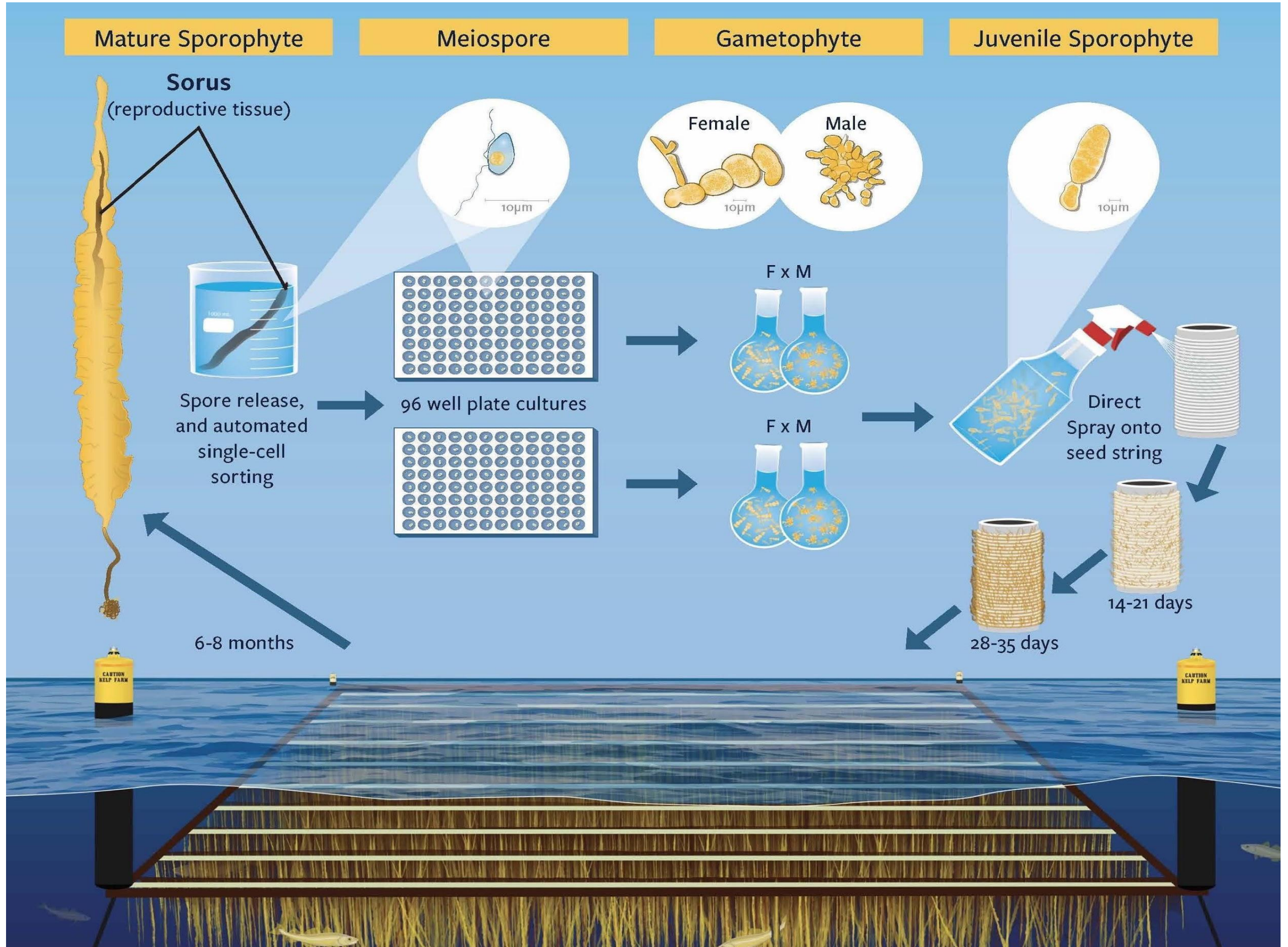
Figure 3.
Change of genetic variance from different breeding schemes over 7 years for a.) 24 or b.) 96 gametophytes per parental sporophyte. The scheme abbreviations are the same as for Figure 2. Each scheme was repeated 100 times and genetic variance shown was the average. The standard error was smaller than the figure symbols and is not shown.

Figure 4
Four views of the simulation results on the final genetic variance and genetic mean. Each view presents the same scatterplot, with each point representing the mean outcome of 100 simulations of one scheme. Each view shows a different obstacle to overcome, with the color of the point determined by the current practice (black) or the improved practice (red). Gray lines connect simulation schemes that are identical except for this change in practice.

bioRxiv preprint doi: https://doi.org/10.1101/2021.01.21.427651; this version posted January 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

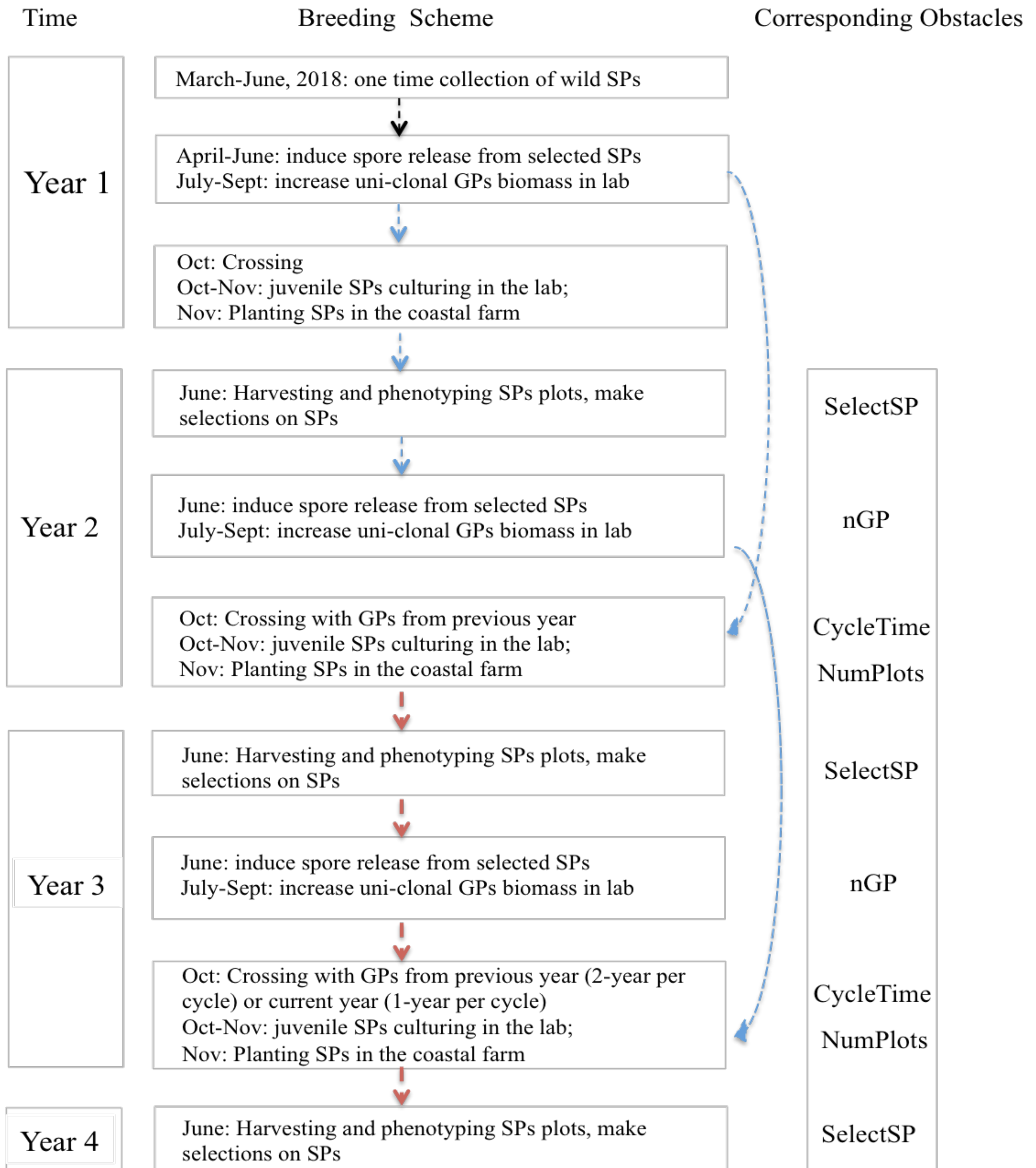
Figure 1

a.



bioRxiv preprint doi: https://doi.org/10.1101/2021.01.21.427651; this version posted January 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

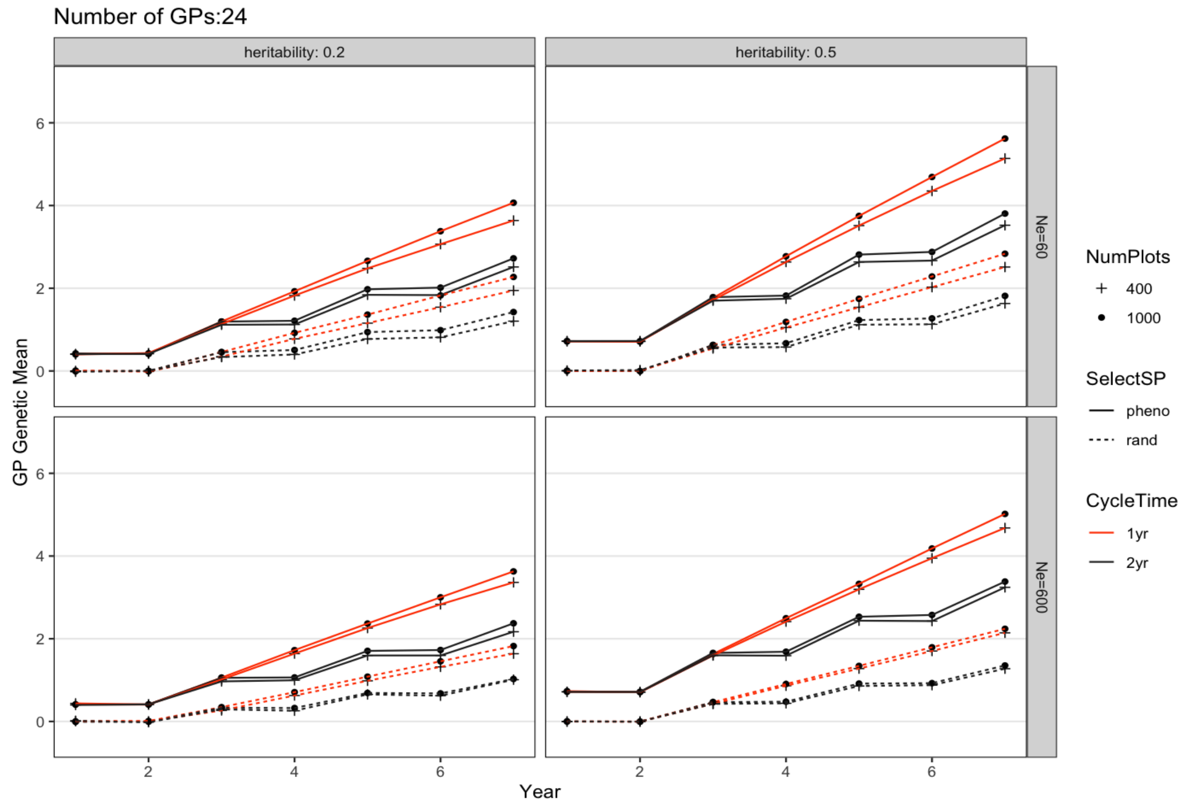
b.



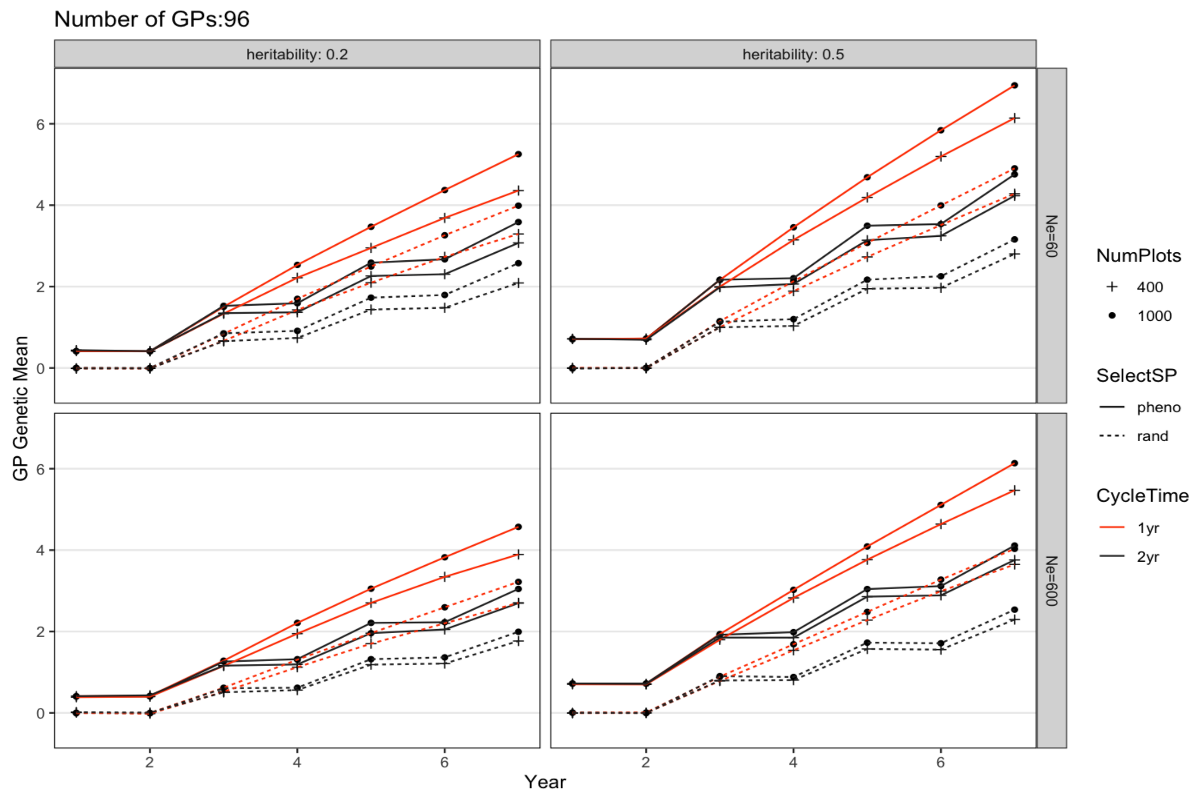
bioRxiv preprint doi: https://doi.org/10.1101/2021.01.21.427651; this version posted January 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Figure 2.

a.



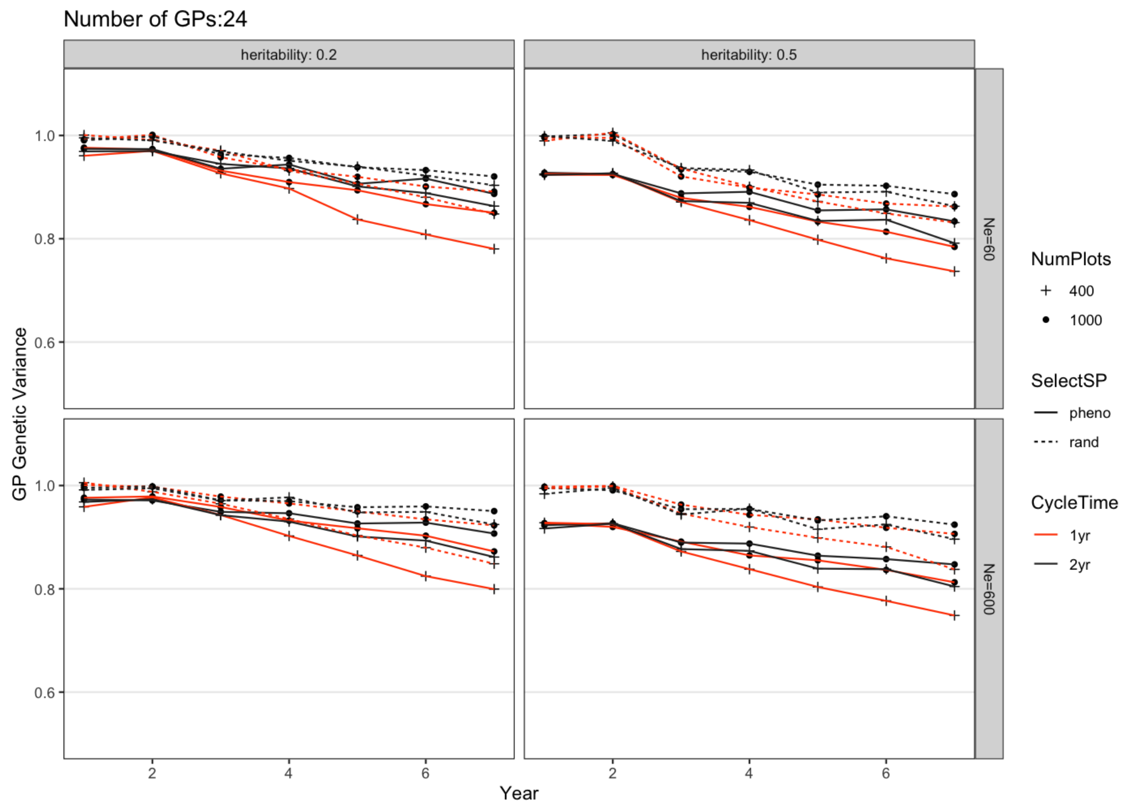
b.



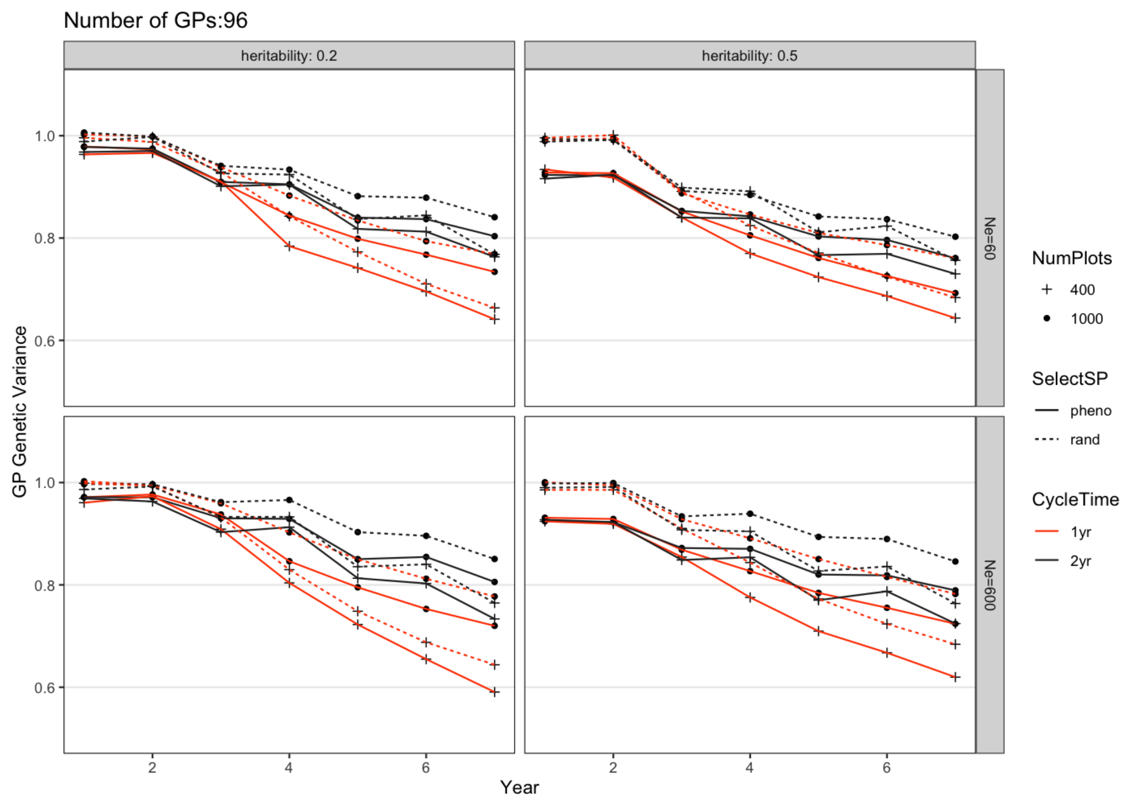
bioRxiv preprint doi: https://doi.org/10.1101/2021.01.21.427651; this version posted January 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Figure 3

a.

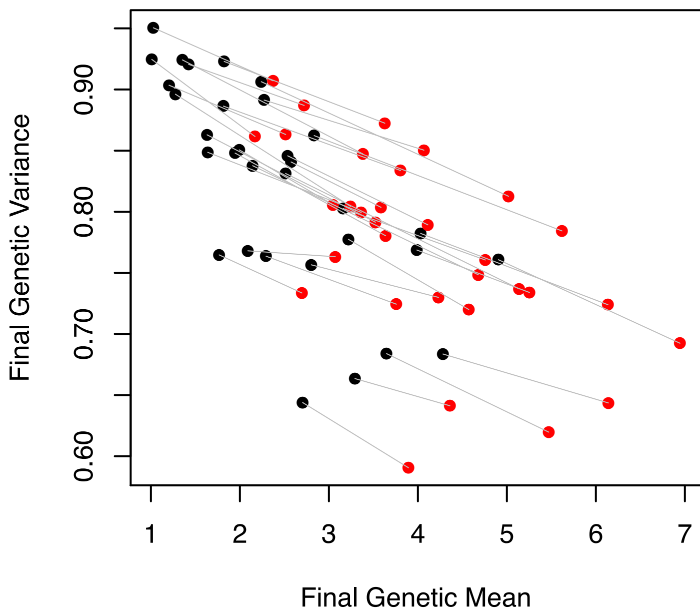


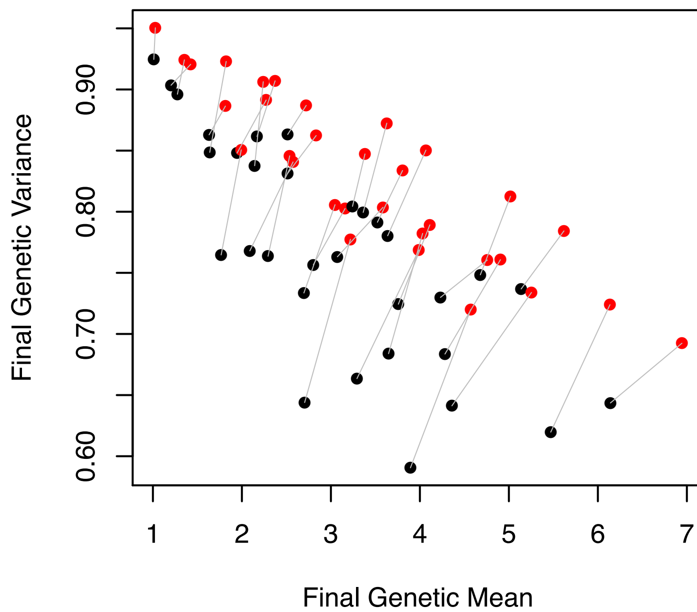
b.

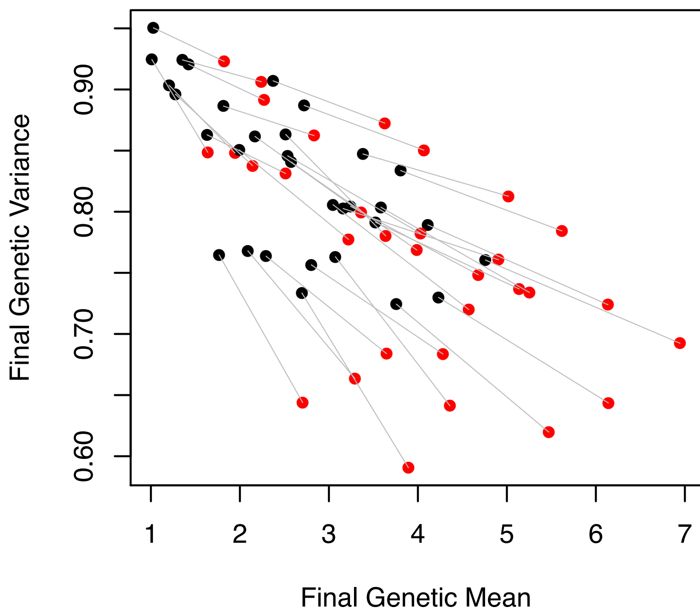


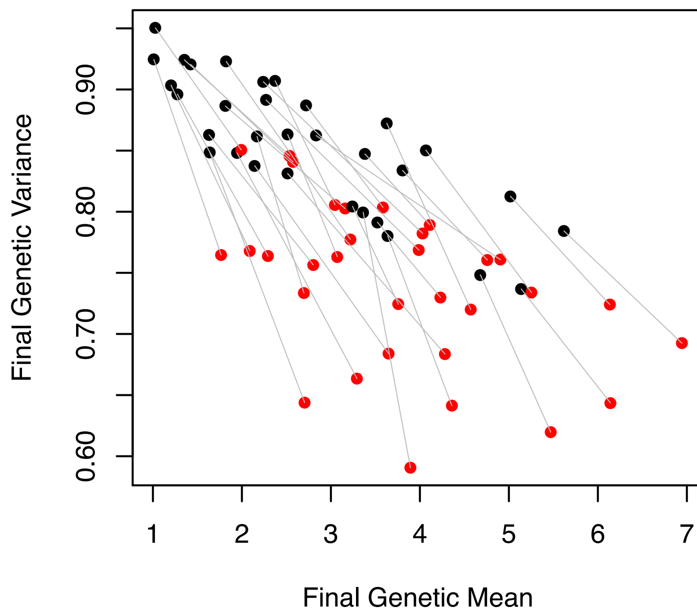
bioRxiv preprint doi: https://doi.org/10.1101/2021.01.21.427651; this version posted January 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Figure 4


Selection on SP
Final Genetic Variance
Final Genetic Mean
0.60
0.70
0.80
0.90
1
2
3
4
5
6
7


Number of Plots Eval.
Final Genetic Variance
Final Genetic Mean
0.60
0.70
0.80
0.90
1
2
3
4
5
6
7


Cycle Time in Years
Final Genetic Variance
Final Genetic Mean
0.60
0.70
0.80
0.90
1
2
3
4
5
6
7


nGP per Parental SP
Final Genetic Variance
Final Genetic Mean
0.60
0.70
0.80
0.90
1
2
3
4
5
6
7