



## Article

# Breeding Alfalfa (*Medicago sativa* L.) Adapted to Subtropical Agroecosystems

Janam P. Acharya <sup>1</sup>, Yolanda Lopez <sup>1</sup>, Beatriz Tome Gouveia <sup>2</sup> , Ivone de Bem Oliveira <sup>3</sup> ,  
Marcio F. R. Resende Jr. <sup>3</sup>, Patricio R. Muñoz <sup>3</sup> and Esteban F. Rios <sup>1,\*</sup>

<sup>1</sup> Agronomy Department, University of Florida, Gainesville, FL 32611, USA; janamacharya@ufl.edu (J.P.A.); ylopezb@ufl.edu (Y.L.)

<sup>2</sup> Departamento de Biologia, Universidade Federal de Lavras, Lavras, Minas Gerais 37200, Brazil; biatgouveia@hotmail.com

<sup>3</sup> Department of Horticultural Sciences, University of Florida, Gainesville, FL 32611, USA; idebem.oliveira@ufl.edu (I.d.B.O.); mresende@ufl.edu (M.F.R.R.J.); p.munoz@ufl.edu (P.R.M.)

\* Correspondence: estebanrios@ufl.edu

Received: 14 April 2020; Accepted: 18 May 2020; Published: 21 May 2020



**Abstract:** Alfalfa is planted in more than 30 million hectares worldwide, but despite its popularity in temperate regions, it is not widely grown in subtropical agroecosystems. It is critical to improve alfalfa for such regions, considering current predictions of global warming and the increasing demands for animal-based products. In this study, we examined the diversity present in subtropical alfalfa germplasm and reported genetic parameters for forage production. An initial screening was performed from 2014 to 2016, evaluating 121 populations from different subtropical origins. Then, a breeding population was created by crossing selected plants, resulting in 145 full-sib and 36 half-sib families, which were planted in a row-column design with augmented representation of three controls ('Bulldog805', 'FL99' and 'UF2015'). Dry matter yield (DMY), canopy height (AH), and percentage blooming (BLOOM) were measured across several harvests. Moderate narrow-sense heritability and high genetic correlations between consecutive harvests were estimated for all traits. The breeding line UF2015 produced higher DMY than FL99 and Bulldog805, and it could be a candidate cultivar release. Several families produced higher DMY than all checks, and they can be utilized to develop high yielding and adapted alfalfa cultivars for subtropical agroecosystems.

**Keywords:** alfalfa; forage; biomass; legume; hay; livestock

## 1. Introduction

Subtropical agroecosystems offer unique abiotic and biotic challenges for agriculture production. Forage breeders aiming at developing cultivars for subtropical regions are challenged to identify cool-season species with sufficient pest/disease resistance [1], while also selecting warm-season species that can extend forage production during transition periods [2–4]. In addition, a primary goal for forage breeders is to select perennial plants that are able to maintain their productivity and stands over several years [5].

Alfalfa (*Medicago sativa* L.) is known as the queen of forages [6] and is the most important forage legume in the world, grown in more than 30 million hectares [7]. It is the fourth most valued crop in the United States (US) after corn, soybean, and wheat, with an estimated value of \$8.8 billion [8]. However, alfalfa production is small in the southeast US, and it has not been recently reported in Florida [9]. The Köppen-Trewartha Climate Classification system has classified Central/North Florida as a Subtropical and Mediterranean climate, and South Florida as a Tropical climate [10]. Moreover, with the current predictions of global warming [11], the tropical savanna climate type is likely to

be expanded throughout Florida in the future [12]. Therefore, it is critical to breed a wide array of forage species adapted to the region's diverse livestock systems, soil types, and climates [13]. Even though alfalfa cultivars were previously released and showed improved adaptation to the state's subtropical agroecosystem (e.g., 'Florida 66' [14], 'Florida 77' [15], and 'Florida 99'), seeds from these cultivars are not commercially available. Therefore, it is imperative to develop new cultivars adapted to subtropical agroecosystems.

Alfalfa is also cultivated in more than four million hectares in South America, with Argentina as the leading producer (3.2 million hectares), followed by Chile and Peru each with (120,000 hectares), and Uruguay (70,000 hectares) [16]. Given its low adaptation to subtropical and tropical climates, alfalfa is only grown in 35 thousand hectares in Brazil [17]. Therefore, the development of adapted cultivars for these challenging environments would offer great potential for alfalfa expansion. However, such a venture would require significant effort in breeding and selection. The investigation of diversity panels and the generation of genetic parameters for agronomic traits would benefit alfalfa breeders in this sense.

Alfalfa is a perennial, outcrossing, and autotetraploid ( $2n = 4 \times = 32$ ) species. Yield improvement in alfalfa is compromised by long selection cycles, tetrasomic inheritance, high inbreeding depression, and considerable genotype and environment interaction for this complex trait [18–20]. Most alfalfa breeding programs use recurrent phenotypic selection to develop and select cultivars with improved yield and tolerance to biotic/abiotic stresses, while new breeding methods and approaches have been implemented lately [21–23]. In the past 60 years, there has been rapid advancement in terms of developing disease resistance, pest management, persistence, and transgenic cultivars in alfalfa [24]. However, there has been little improvement related to crop yield, and alfalfa yields have stagnated in the recent years [25]. Narrow-sense heritability ( $h^2$ ) of 0.15 and broad-sense heritability ( $H^2$ ) of 0.30 were previously reported for yield [25,26]. However, most of these yield studies in alfalfa were performed in short-term experiments or few harvests (three to four harvests per year) [7,27], with the exception of de Assis [17], who used 23 harvests for the rainy season in the summer, and eight harvests for dry season in the winter, and conducted the experiment for four years. Reliable estimates of  $h^2$  and other genetic parameters are essential to explain optimal selection schemes [28]. However, information regarding genetic parameters in alfalfa in tropical and subtropical conditions is still meager [7].

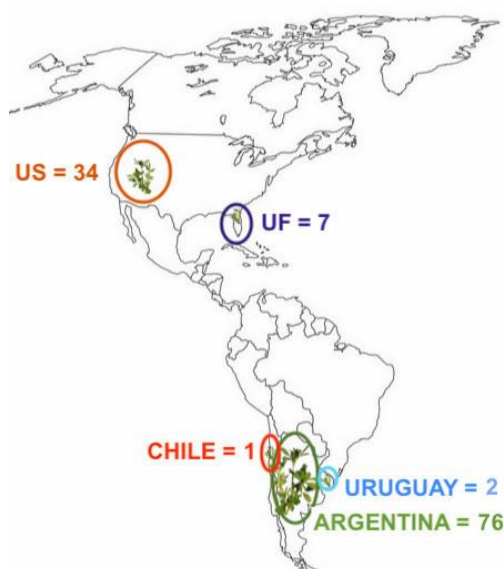
The use of best linear unbiased prediction (BLUP) is an established method to predict genetic and breeding values in animals, and has been extensively used for plant species helping guiding decisions towards breeding. The use of BLUP can generate accurate predictions of genotypic values even for unbalanced experimental designs, and BLUPs are also used to estimate genetic correlations among performance of the same genotypes in different environments [29]. Here, we aimed to generate a reference breeding population and used BLUP to generate information concerning genetic parameters for important agronomic traits in alfalfa, as well as to obtain genotype rankings, allowing selection of alfalfa genotypes adapted to tropical/subtropical conditions. Our objectives were: i) to screen alfalfa germplasm for biomass production in Florida; ii) to develop and phenotype a reference breeding population for agronomic traits at the family level; and iii) to estimate variance components and genetic parameters and use them to obtain rankings based on genotypic values, to select families with improved and desirable agronomic traits.

## 2. Materials and Methods

### 2.1. Initial Germplasm Screening

A screening trial consisting of 121 alfalfa populations (including breeding lines and cultivars) obtained from five tropical/subtropical germplasm pools: Argentina (76), Chile (1), Uruguay (2), United States, Florida (7), and other sources (34) (Figure 1) was conducted from October 2014 to April 2016. Fall dormancy groups ranged from 5 to 10 in these populations. The experiment was established as a randomized complete block design, with two replicates at the Plant Science Research

and Experimental Unit (PSREU), Citra, FL, USA. The soil at the site was Chipley sand soil (Thermic coated Aquic Quartzipsamments) and historical weather data was obtained from the Florida Automated Weather Network (<https://fawn.ifas.ufl.edu>). Experimental units were planted by broadcasting the equivalent to 22 kg/ha of seed per plot (1 m × 1 m) on 27 October 2014. Treflan (1.7 L·ha<sup>-1</sup>) was applied a week prior to planting as a pre-emergence herbicide. Weeds were manually controlled after planting, and irrigation was supplemented as needed to promote establishment and regrowth after each harvest. Experimental units received 67.25 kg·K<sub>2</sub>O·ha<sup>-1</sup> as Muriate of Potash and 1.12 kg·ha<sup>-1</sup> boron after each harvest. Dry matter yield (DMY) was assessed across ten harvests during spring/summer/fall in 2015 and spring of 2016. A total of 33 populations were selected, based on high DMY and persistence across all harvests. In order to generate a reference breeding population by performing controlled crosses, individual vigorous plants were selected within each of these 33 populations to be used as parents.



**Figure 1.** Germplasm sources and number of populations from each source for the initial alfalfa screening conducted in Citra, FL.

## 2.2. Development of the Reference Breeding Population

Crosses of the selected individual plants were performed following a factorial mating design. Male parents consisted of six plants with known adaptation to Florida (four plants from the UF breeding program, and two commercial cultivars: ‘Bulldog805’ and ‘AmeriStand915’), and a set of 27 selected plants were used as females. Fall dormancy ranged from 8 to 10 among male parents, and from 6 to 10 among female parents. All crosses were conducted in controlled conditions in the Forage Breeding and Genetics Lab greenhouse, at University of Florida (Gainesville, FL, USA). Seeds from each full-sib and half-sib family were harvested, threshed individually, planted in 72-cell styrofoam trays in August 2017, and maintained in the greenhouse until transplanting in November 2017.

## 2.3. Experimental Design and Field Management

The field experiment to evaluate the reference population was conducted from November 2017 to March 2019 at PSREU, Citra, FL. Soil samples were taken during land preparation and tested at the Extension Soil Testing Lab at the University of Florida. Soil pH was 6.8, and soil P, K and Mg were high, medium, and high respectively. The experiment was established in a row-column design [30] with augmented representation of three controls (i.e., cv. Bulldog805, cv. ‘FL99’, and advanced breeding line UF2015). Each experimental unit (1.82 m × 1.82 m) consisted of eight rows spaced at 22.8 cm. Twenty seedlings per plot were manually transplanted in the middle rows. The remaining six rows

(three on each side) were seeded (drill) with cv. Bulldog805 at  $20 \text{ kg} \cdot \text{ha}^{-1}$ , to serve as border for each plot. The field was fertilized after each harvest with  $67.25 \text{ kg} \cdot \text{K}_2\text{O} \cdot \text{ha}^{-1}$ , using Muriate of Potash, and with Boron at the rate of  $1.12 \text{ kg} \cdot \text{ha}^{-1}$ . The experiment was conducted in rainfed conditions, weeding was done manually for broad-leaf weeds after each harvest, and herbicide Clethodim (Select,  $70.76 \text{ g AI/L-1}$ ; Valent USA Corporation, Walnut Creek, CA, USA) was applied at the rate of  $1.05 \text{ kg} \cdot \text{ha}^{-1}$  after harvesting to control grasses.

#### 2.4. Data Collection

The experimental units were harvested for the first time in April 2018, when the reference breeding line UF2015 reached 10% blooming. The same threshold was applied for harvests throughout the study. Data collection was initiated by mowing border rows at 6 cm stubble height using a flail mower, and biomass was removed from the area. Later, traits were measured at the plot/family level. The average canopy height (AH, cm) was assessed using a ruler placed in the center of each plot, and percentage blooming (BLOOM) was visually rated (0 to 100%). Experimental units were manually harvested, and total fresh weight was recorded. Approximately 500 g of vegetative material was collected for each plot and dried at  $65^\circ \text{C}$  for 7 days, to determine dry matter yield (DMY). The cycle for data collection (mowing borders, AH, and DMY) was repeated 11 times from April 2018 until May 2019.

#### 2.5. Statistical Analyses

First, analyses were conducted by individual harvest, and then a multi-harvest model was fitted. All analyses were performed using linear mixed models in ASReml-R version 4.0 [31], implemented in R [32]. The pedigree of all the families were known, and the additive pedigree relationship matrix was constructed using AGHmatrix 0.0.4 in R [33], considering the tetraploid option. Single harvest analysis was performed using the model:

$$y = \mu + Xt + Z1r + Z2c + Z3f + e, \quad (1)$$

where,  $\mu$  is the overall mean;  $X$  and  $Z$  represent the incidence matrices for fixed and random effects, respectively;  $t$  is the fixed effect for the checks;  $r$  is the random effect of row,  $r \sim \text{MVN}(0, I\sigma_r^2)$  and  $\sigma_r^2$  is the variance of row;  $c$  is the random effect of column,  $c \sim \text{MVN}(0, I\sigma_c^2)$  and  $\sigma_c^2$  is the variance of column;  $f$  is the random vector of family,  $f \sim \text{MVN}(0, A\sigma_f^2)$  and  $\sigma_f^2$  is the variance of family;  $e$  is the random vector of error,  $e \sim \text{MVN}(0, I\sigma_e^2)$ ; and  $\sigma_e^2$  is the variance of residual.  $A$  is the relationship (kinship) matrix and  $I$  is an identity matrix of its proper size.

Then, a model using harvest as repeated measures was fitted to obtain a single genotypic value across harvests, where the following model was used for the multi-harvest analysis for single traits:

$$y = \mu + X1h + X2t + Z1r + Z2c + Z3fh + e \quad (2)$$

where,  $\mu$  is the overall mean;  $X$  and  $Z$  represent the incidence matrices for fixed and random effects, respectively;  $h$  is the fixed effect of harvest;  $t$  is the fixed effect for the controls;  $r$  is the random effect of row,  $r \sim \text{MVN}(0, I\sigma_r^2)$  and  $\sigma_r^2$  is the variance of row;  $c$  is the random effect of column,  $c \sim \text{MVN}(0, I\sigma_c^2)$  and  $\sigma_c^2$  is the variance of column;  $fh$  is the random vector of family within each harvest,  $fh \sim \text{MVN}(0, G \otimes A)$ ;  $e$  is the random vector of error within each harvest,  $e \sim \text{MVN}(0, R \otimes I)$ ; and  $\sigma_e^2$  is the variance of residual.  $R$  is the residual variance/covariance matrix,  $G$  is the genetic variance/covariance matrix,  $A$  is the relationship (kinship) matrix, and  $I$  is an identity matrix of its proper size. The Kronecker product is denoted by  $\otimes$ .

The modeling of the matrix of residual ( $R$ ) and genetic ( $G$ ) variance/covariance matrix was performed for each trait in the multi-harvest model, Equation (2). The choice of the matrices was done sequentially [34,35], first identifying the best structure for  $R$ , and then for  $G$ , considering  $R$  previously selected. The best structure for  $R$  and  $G$  was selected based on the Bayesian Information criterion (BIC),

as proposed by Schwarz [36]. Variance component estimates from Equations (1) and (2) were used to calculate family-based narrow-sense heritability ( $h^2$ ) for the breeding population as:

$$h^2 = \sigma_f^2 / \sigma_P^2 = \sigma_f^2 / (\sigma_c^2 + \sigma_r^2 + \sigma_f^2 + \sigma_e^2), \quad (3)$$

where,  $\sigma_P^2$  is the phenotypic variance. In the multi-harvest analysis, the average of family and residual variance of the harvests were used to estimate the  $h^2$ .

The Spearman correlation of family performance across harvests was calculated using the BLUPs estimated for families using the single harvest analysis, using the package *agricolae* [37] in R [32]. The significance of correlation estimates was verified by Spearman's  $\rho$  test, using the package *Stats* in R [32]. Genetic correlations among traits (*Type-A* correlation) were estimated using a bivariate model, where  $y_i$  was partitioned for a 2-trait analysis  $y_1$  and  $y_2$ . The 2-trait (bivariate) model fitted below:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \mu_i + X_1 t + X_2 h(t) + X_3 d(t) + Z_1 r(t) + Z_2 c(t) + Z_3 f(t) + Z_4 fh(t) + Z_5 rh(t) + Z_6 ch(t) + e, \quad (4)$$

where  $y_1$  and  $y_2$  are the data vectors for the two traits of interest;  $\mu_i$  is the overall population mean for each trait;  $h(t)$  and  $d(t)$  are, respectively, the vectors of the fixed effects of harvests and controls within trait;  $r(t)$  is the random vector of row effects within traits, with  $r(t) \sim MVN(0, I \otimes B)$ ;  $c(t)$  is the random vector of column effect within traits, with  $c(t) \sim MVN(0, I \otimes C)$ ;  $f(t)$  is the random vector of family effects within traits, with  $f(t) \sim MVN(0, I \otimes F)$ ;  $fh(t)$  is the random vector of family  $\times$  harvest interaction effects within traits, with  $fh(t) \sim MVN(0, I \otimes H)$ ;  $rh(t)$  is the random vector of row  $\times$  harvest interaction effects within traits, with  $rh(t) \sim MVN(0, I \otimes J)$ ;  $ch(t)$  is the random vector of column  $\times$  harvest interaction effects within traits, with  $ch(t) \sim MVN(0, I \otimes K)$ ; and  $e$  is the random vector of errors, with  $e \sim MVN(0, I \otimes R)$ .  $B$ ,  $C$ ,  $F$ ,  $H$ ,  $J$ , and  $K$  are compound symmetry heterogeneous (CSH)  $2 \times 2$  variance-covariance matrices between traits defined by a single trait-to-trait *type-A* genetic correlation term ( $r_A$ ), and a unique  $j$ th variance term,  $\sigma_{gtj}^2$  for each trait.  $R$  is an unstructured  $2 \times 2$  residual variance-covariance matrices with a different variance for each trait ( $\sigma_{ej}^2$ ) and a covariance between traits ( $\sigma_{eij}$ ).  $I$  is an identity matrix of its proper size and the Kronecker product is denoted by  $\otimes$ . The genetic correlations ( $r_g$ ) for each pair of traits (*type-A*) were estimated directly from the genetic variance-covariance matrices from the models described above, where values close to 1 or -1 indicate a strong association between genetic values from a pair of traits.

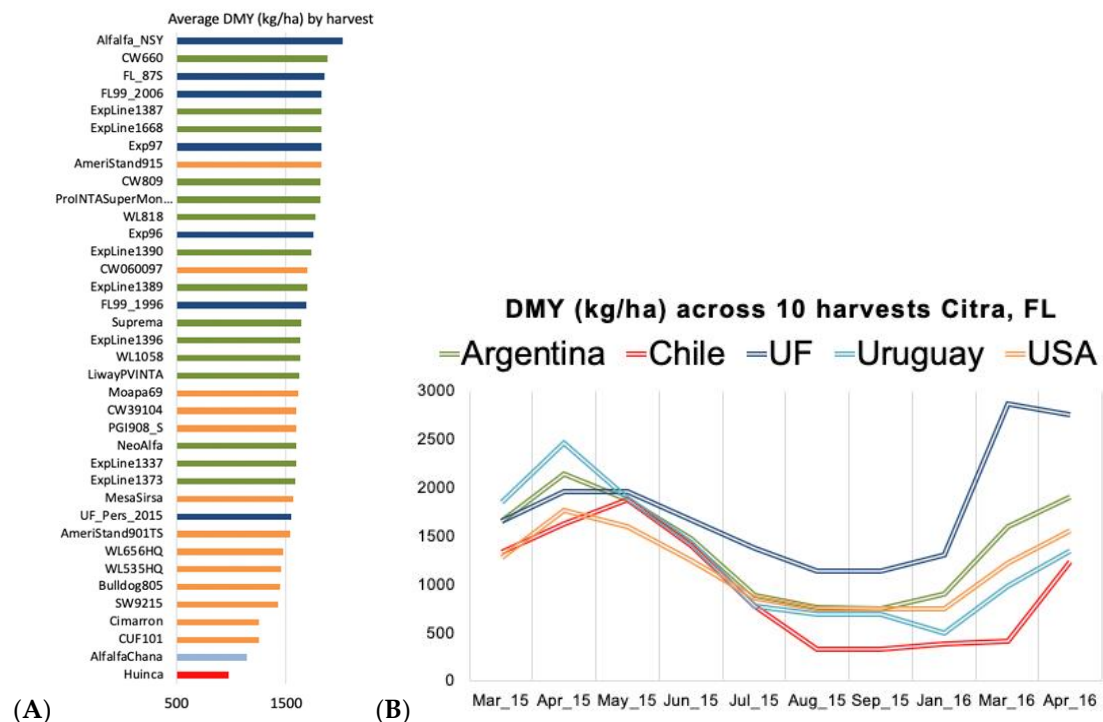
Principal component analysis (PCA) was performed with the *prcomp* function, as implemented in R [32], calculated through a correlation matrix using the genotypic values obtained with the multi-harvest model. The PCA results were visually plotted using the package *ggbiplot* in R [38]. For all other analysis, graphs were created using *ggplot2* [38].

### 3. Results

#### 3.1. Initial Germplasm Screening and Mating Design

The initial germplasm screening of 121 populations showed significant phenotypic diversity for DMY after 10 harvests, and DMY per harvest ranged from 450 to 2000 kg DMY/ha (Figure 2). Several populations produced average DMY per harvest lower than 1000 kg DMY/ha, and were not selected as parents for subsequent crosses. Seven populations from the UF Agronomy Department Alfalfa Breeding Program exhibited high DMY, particularly for harvests performed after the summer 2015 (Figure 2A,B). Fifteen populations from Argentina and thirteen from other US sources also produced more than 1000 kg DMY/ha per harvest on average (Figure 2A,B). The average DMY among populations for each germplasm origin (Figure 2B) showed that these populations produced more dry matter in the first spring after planting (Spring 2015), as compared to summer 2015. The UF germplasms produced higher DMY from July 2015 to the end of the study. All populations exhibited an increase in DMY in late-winter/early-spring of the second year (Figure 2B).



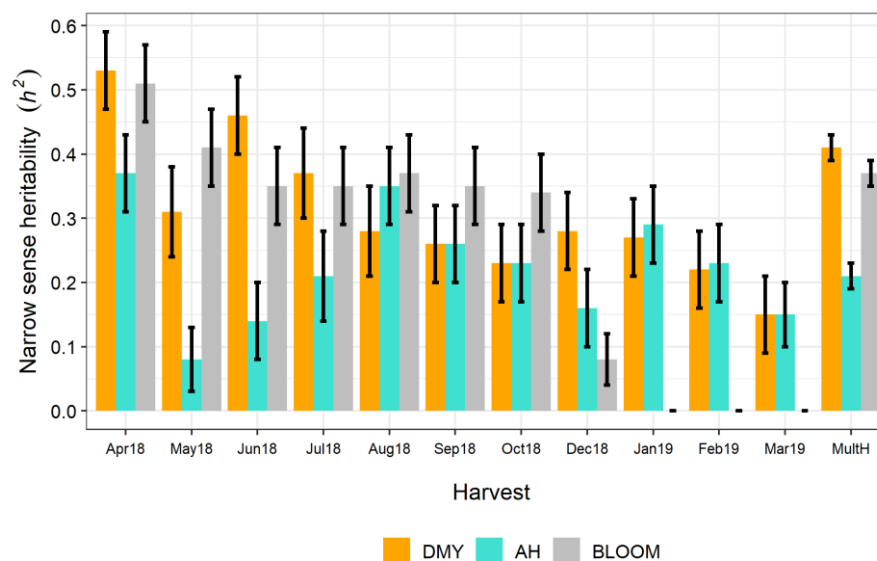


**Figure 2.** Alfalfa germplasm screening for dry matter yield (DMY, kg/ha) conducted in Citra, FL using 121 alfalfa populations from five different sources (blue: University of Florida, green: Argentina, light blue: Uruguay, red: Chile, and orange: USA states other than Florida). (A) Average DMY across 10 harvests for the populations exhibiting higher DMY, (B) average DMY among germplasm source.

The ranking for average DMY per harvest was used to select populations exhibiting adaptation to Florida, and the populations with higher DMY are shown in Figure 2A. A single plant per population was selected based on vigor, and cuttings were made in the summer 2016. Fall dormancy ranged from eight to 10 among male parents, and from six to 10 among females. A factorial mating design was used to create all possible full-sib combinations; however, some crosses did not produce enough viable seed, and were not included in further trials.

### 3.2. Variance Components and Genotypic Values

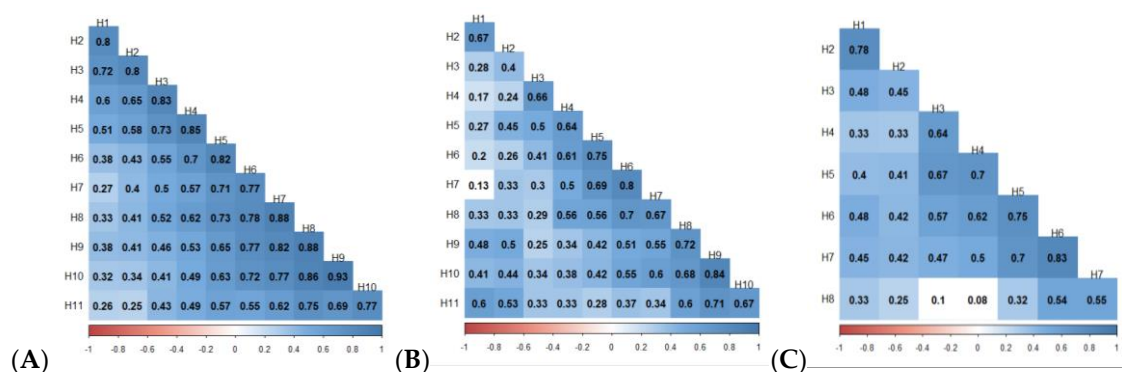
The modeling of the matrix of residual (R) and genetic (G) variance and covariance were performed for each trait and harvest, identifying the best structure for R first, and then for G (Supplemental Table S1). The diagonal matrix and autoregressive of order one heterogeneous structure provided the lowest BIC value for R and G matrix for all traits, respectively. The alfalfa reference breeding population exhibited significant genetic variation ( $p < 0.001$ ) across eleven harvests for all traits (Supplemental Tables S4–S6). Low to moderate  $h^2$  (0.08 to 0.53) was estimated for all traits in each harvest (0.15 to 0.53 for DMY, 0.08 to 0.37 for AH, 0.08 to 0.51 for BLOOM). All three traits had the highest  $h^2$  estimates in the first harvest, whereas lower estimates were obtained in later harvests (Supplemental Tables S2–S4 and Figure 3). However, moderate values were observed for the multi-harvest analysis for each trait. The estimate of  $h^2$  was 0.41 (DMY), 0.21 (AH), and 0.37 (BLOOM) across all the harvests, using the multi-harvest model (Figure 3).



**Figure 3.** Narrow-sense heritability ( $h^2$ ) and standard error (black error bar) for dry matter yield (orange), canopy height (green), and percent blooming at harvest (gray) in a single harvest (denoted by month and year), and in multiple-harvest (MultH) analysis in alfalfa.

### 3.3. Spearman Correlation Between Harvests, and Type-A Genetic Correlation Between Traits

Spearman correlations ranged from 0.26 to 0.93 for DMY (Figure 4A), 0.13 to 0.84 for AH (Figure 4B), and 0.08 to 0.84 for BLOOM (Figure 4C). All traits exhibited higher correlation coefficients between consecutive harvests, whereas moderate values were obtained among distant harvests. All traits exhibited positive correlation with each other. The highest type-A genetic correlation ( $r = 0.85$ ) was estimated between AH and DMY; followed  $r = 0.57$  between AH and BLOOM, and  $r = 0.42$  between DMY and BLOOM.

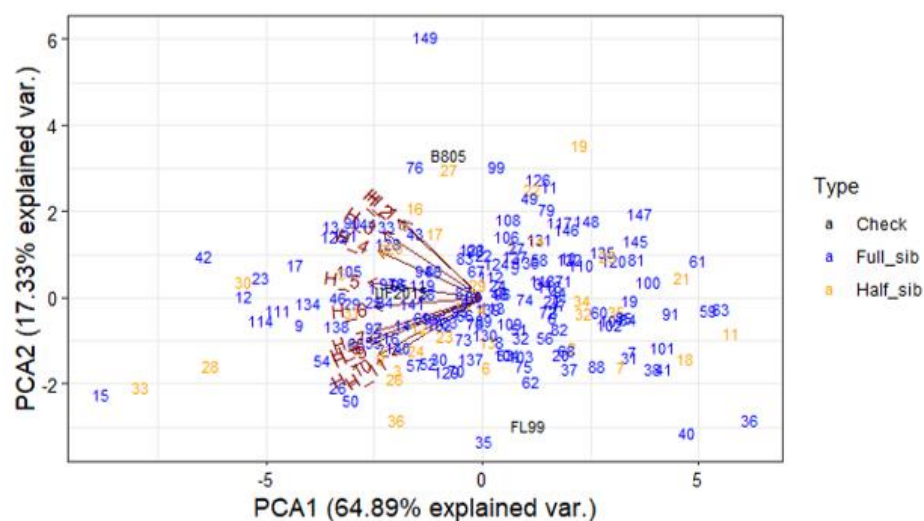


**Figure 4.** Spearman correlations across harvests using best linear unbiased predictions (BLUPs) from single harvest analysis for dry matter yield (A), average canopy height (B) and percent bloom (C) measured in 145 full-sib, 33 half-sib alfalfa breeding populations and three controls (cv. Bulldog805, cv. FL99, and advanced breeding line UF2015) across 11 harvests (eight for percentage blooming (BLOOM)) in Citra, FL. Color-coded panel indicates the magnitude of the correlation coefficient, and white panels indicate non-significant correlations by Spearman's rho test at 5% probability.

### 3.4. Principal Component Analysis for DMY

Principal Component Analysis showed that the first two components (PCA1 and PCA2) accounted for 81% of the variation in DMY genotypic values across the eleven harvests (Figure 5). The first component (PCA1) explained 64% of the variation, and the second component (PCA2) explained 17% of the variation. Harvest five and harvest six contributed most towards the variation in the first

component, while harvest one, two, ten, and eleven contributed more towards PCA2. No distinct clustering was observed for half- and full-sib families, and a broad range of variation was observed for DMY across harvests. For the controls used in this study, UF2015 produced the highest DMY across all harvests. Significant phenotypic variability was present in the reference breeding population for DMY, with some families producing more DMY at the beginning of the experiment, while their production was significantly reduced after the summer harvests (Figure 5). Some families exhibited improved persistence and their DMY was high throughout the study. Specifically, full-sib family 149F produced the highest DMY during the initial harvests; however, families 15F, 33H, 28H, 42F exhibited higher DMY in later harvests. Several half- and full-sib families produced lower yields than the three checks across all harvests (Figure 5).



**Figure 5.** Principal component analysis (PCA) for DMY across eleven harvests (vectors), using the family predicted values on each harvest (loadings). Color-coded loadings represent the three alfalfa checks (black), half-sib families (orange), full-sib families (blue).

#### 4. Discussion

The projected increase in the global population to 9.2 billion in 2050 demands an increase in agriculture production, including animal-based products [39]. A projected increase of 200 M tons of beef and other livestock products is necessary to meet the future demand [39], particularly in subtropical/tropical regions such as Africa [40] and South America. This unprecedented increase in animal-based products will rely on adequate animal feed supplies. Alfalfa is the most commonly grown forage, and its international trade for hay reached 8.3 million metric tons in 2017, for a total value of USD 2.3 billion, and the demand for alfalfa hay will continue to increase [16]. Despite being the most important forage legume crop globally, limited resources have been dedicated towards improving alfalfa adaptation to subtropical and tropical regions.

The Alfalfa Research and Development Latin American Platform [16] promotes the cultivation of alfalfa in South America, and they have made significant efforts to breed alfalfa adapted to subtropical [41] and tropical [17] agroecosystems. In the US, improved varieties were developed for the southeastern region [5,42], but most of these cultivars did not persist well under the hot and humid conditions prevalent in Florida (Figure 2A). Three well-adapted cultivars ('Florida 66', 'Florida 77', and 'Florida 99') previously developed and released by the UF Agronomy Department [14,15] are not commercially available anymore. Therefore, there is a growing need to develop germplasm well-adapted to subtropical agroecosystems and release cultivars with improved DMY and persistence. The goal of this study was to create a reference breeding population using parental lines adapted to subtropical and tropical environments, in order to develop suitable cultivars for tropical and subtropical climates, such as Florida. The phenotypic diversity and genetic parameters presented in this study



exhibited the existence of adapted germplasm, and the potential for developing and releasing cultivars in the near future.

#### 4.1. Initial Germplasm Screening and Mating Design

The introgression of germplasm from international sources proved to be a successful approach to increase genetic diversity for DMY and persistence in the UF program. The seven breeding populations from Florida exhibited higher DMY after the fifth harvest in the summer 2015, until the end of the study, showing the importance of developing locally-adapted germplasm. These seven populations were at the top of the ranking for average DMY across 10 harvests (Figure 2A). Similarly, the locally developed cultivars “Florida66” and “Florida77” showed improved persistence, and consequently higher yields in the second production year, compared to other cultivars in Florida [14,15]. Nevertheless, several populations from Argentina and other US sources also exhibited high DMY per harvest, even after the fifth harvest (Figure 2A).

Selection based on DMY across harvests resulted in 33 populations exhibiting improved yield in Florida, and a single plant per population was used as a parent in crosses to develop a breeding population. The mating design allowed the utilization of at least one male parent with proven adaptation to the Southeastern US, and particularly to Florida. Due to the diverse geographical origin among germplasm sources, it was expected that these parental lines would be genetically distant, and therefore would provide novel allele combinations to broaden the genetic base in the UF alfalfa breeding program. As a future study, molecular markers will provide more detailed data regarding the genetic diversity present in the parental lines and their progeny.

#### 4.2. Variance Components and Genotypic Values for the Reference Breeding Population

Results from this study indicated that significant genetic variation existed for the three traits in the reference breeding population. Phenotypic characterization showed a wide range of continuous variation for the three agronomic traits, and the presence of a wide range of variation among families, which would warrant the improvement of germplasm adapted to Florida. The wide range of phenotypic variation can be used to guide plant selections for future cultivar releases, as well as to use this germplasm for novel breeding approaches such as genomic selection and marker assisted breeding.

Family-based  $h^2$  for DMY for the multi-harvest model ( $0.30 \pm 0.02$ ) was higher than previous reports [7,26]. Bowley and Christie [26] conducted an experiment based on 90 genotypes and three harvests; while Annicchiarico [7] used 125 genotypes, representative of Italian germplasm, which were evaluated for 12 harvests. Riday and Brummer [27] reported DMY  $h^2$  of 0.12, 0.27, and 0.34 for three harvests. In our study, the lowest DMY  $h^2$  (0.15) was estimated for the last harvest, whereas the highest ( $h^2 = 0.53$ ) was estimated in the first harvest. The higher values obtained for heritability in our study may be associated with the diverse population produced by the combination of germplasm from different breeding programs, which generated a high degree of genetic variability. Nevertheless, we had some families that performed better in the earlier harvest and their yield declined rapidly, perhaps due to lack of persistence. Therefore, selection in alfalfa for DMY in early harvests (prior to first summer after planting) may result in non-persistent germplasm, which would limit long-term production in subtropical regions.

The  $h^2$  of AH was 0.21 ( $\pm 0.02$ ) for the multi-harvest model, which was lower than previous reports by Riday and Brummer [27]. Riday and Brummer [27] reported  $H^2$  of 0.40, 0.46, and 0.41 for three harvests respectively. However, a higher number of harvests were evaluated in our study, and the  $h^2$  was low to moderate for each harvest. Most of the agronomic traits evaluated are quantitative in nature, and are governed by many genes and have a low heritability suggesting the requirement of more time, effort, and labor for the genetic improvement of these traits. BLOOM being an oligogenic trait generally has high heritability but, in this case, it was moderate ( $0.37 \pm 0.02$ ).

#### 4.3. Type-B ( $r_{GxH}$ ) Genetic Correlation Between Families and Harvest

High  $r_{GxH}$  were estimated between consecutive harvests for all the three traits, suggesting that genotypic performance is stable in the short-term, and that fewer harvests could be performed to phenotype these traits. The high  $r_{GxH}$  indicates that the rankings will not change drastically in each of the consecutive harvests. Therefore, reducing the number of harvests, and consequently labor, could allow breeders to increase their breeding population size, and consequently increase genetic gains. In addition, these results demonstrate the need for performing several harvests in alfalfa when selecting for quantitative traits, particularly for harvests performed after the first summer.

#### 4.4. Type-A ( $r_G$ ) Genetic Correlation Between Traits

The *type-A* genetic correlation measures the proportion of variance that two traits share, due to genetic causes. It is an estimate of the degree of pleiotropy or causal overlap, and it is calculated by extracting information from the additive genetic variances and covariances between the traits [43]. In alfalfa, genetic correlations are based on the genetic variance of the traits and covariance between the traits [44], and the additive genetic correlations are based on half-sib family variances and covariances that includes a dominance variance component [45]. The knowledge of  $r_G$  is a very important genetic parameter that can assist selection. High positive  $r_G$  suggests that indirect selection for both traits could be achieved by measuring only one trait, reducing the amount of resources and time in phenotyping, considering a moderate/high heritability for both traits [46]. High  $r_G$  was estimated between AH and DMY, indicating that the taller plants also produced higher DMY. This might be due to the presence of more biomass due to thickening of stems, and a greater number of leaves in those stems. Since there was a high  $r_G$  between AH and DMY, breeders can use AH for selecting higher DMY when large nurseries are to be screened [46]. The AH can also be accurately phenotyped using unmanned aerial vehicles (UAVs), thus making the utilization of high-throughput phenotyping possible in alfalfa breeding programs [47]. There was a positive  $r_G$  correlation between BLOOM and AH, suggesting that taller plants produced flowers earlier. Positive but low correlation  $r_G$  was also observed between BLOOM and DMY. This lower correlation might be due to the harvesting of all the plots when the flowering of UF2015 reached 10%. During this time, some plots in the experiment did not flower, whereas some plots had 100% flowering. This elucidates that the BLOOM ratings used in this protocol should not be used for the indirect measurement of DMY.

#### 4.5. Principal Component Analysis

The PCA analyses also revealed a high family by harvest interaction for distant harvests, and significant genotypic variation across families for DMY. Some families had high DMY in the first four harvests (i.e., 149F, 76F, and the check Bulldog 805), while their DMY decreased in later cuts. Other full- and half-sib families exhibited high DMY across all harvests (i.e., 42F, 23F, 30H, and 12F). A third group of families exhibited higher DMY, particularly at later harvests (15F, 33H, and 28H). The PCA for DMY showed no distinct pattern for full and half-sib families, suggesting that improvements in alfalfa DMY can be performed using half-sib families. The resources and labor in creating full-sibs can be significantly reduced by allowing selected parents to cross under an open-pollination scheme, in isolated environments [41].

Harvest five and six contributed more in explaining the variation in the first component, indicating that efficient selection can be done for DMY in the fifth or sixth harvest, which coincides with harvests performed after the first summer in Florida. Weather conditions (high temperature, rainfall, water table depth) during summer in subtropical environments can be very deleterious to alfalfa stands [48]. These abiotic stresses, coupled with several biotic challenges present in subtropical/tropical environments, impact DMY and persistence in temperate legumes [1,5,17].

## 5. Conclusions

A reference alfalfa breeding population was created using a broad germplasm pool obtained from different subtropical pools, particularly Florida, Argentina, and other US sources. Breeding populations from the UF alfalfa breeding program produced higher DMY and showed improved persistence in Florida; nevertheless, populations from Argentina and other US programs also exhibited high DMY. Full and half-sib families were created and phenotyped for DMY, AH, and BLOOM across 11 harvests. These traits exhibited a wide range of genetic variability. Estimates for  $r_{G \times H}$  indicated that fewer harvests can be performed to phenotype breeding populations and successfully select improved families. The breeding line UF2015 was the best check producing higher DMY across harvests, and it could be a candidate line for cultivar release in the near future. Several families produced high DMY in early harvests, but their DMY decreased after the first summer. Several full and half-sib families produced more DMY than all the checks, exhibiting an improvement in this germplasm. High  $r_G$  correlations indicated the potential for indirect selection between these traits, particularly for AH and DMY. Novel families developed during this experiment showed higher DMY and AH, and they can be utilized to develop high yielding and adapted alfalfa cultivars in Florida, and for other subtropical agroecosystems.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/2073-4395/10/5/742/s1>, Table S1: Bayesian information criterion (BIC) and number of parameters (NP) for different structures of residual (R) and genetic (G) variance/covariance matrix in the multi-harvest model for dry matter yield (DMY), average height (AH) and percent blooming at harvest (BLOOM), Table S2: Variance components, likelihood test, narrow-sense heritability ( $h^2$ ) and standard error (SE) of  $h^2$  for dry matter yield (DMY), Table S3: Variance components, likelihood test, narrow-sense heritability ( $h^2$ ) and standard error of  $h^2$  for average height (AH), Table S4: Variance components, likelihood test, narrow-sense heritability ( $h^2$ ) and standard error of  $h^2$  for percent blooming at harvest (BLOOM).

**Author Contributions:** Conceptualization, P.R.M., Y.L. and E.F.R.; Data curation, J.P.A., Y.L., B.T.G. and I.d.B.O.; Formal analysis, J.P.A., B.T.G., I.d.B.O. and M.F.R.R.J.; Funding acquisition, P.R.M.; Investigation, J.P.A., Y.L., B.T.G., P.R.M. and E.R.; Methodology, J.P.A., Y.L., P.R.M. and E.F.R.; Project administration, P.R.M. and E.F.R.; Resources, P.R.M. and E.F.R.; Software, B.T.G. and I.d.B.O.; Supervision, M.F.R.R.J., P.R.M. and E.F.R.; Visualization, J.P.A. and E.F.R.; Writing—original draft, J.P.A. and E.F.R.; Writing—review & editing, J.P.A., Y.L., B.T.G., I.d.B.O., M.F.R.R.J., P.R.M. and E.F.R. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by Agriculture and Food Research Initiative grant number 2014-67013-22418 to P.R. Munoz, J.W. Olmstead and J.B. Endelman from the USDA National Institute of Food and Agriculture.

**Acknowledgments:** The authors would like to thank Daniel Basigalup from INTA-Manfredi, Cordoba, Argentina, and Ing. Agr. Fernando Ortega from INIA-Chile for kindly providing such valuable germplasm. The authors greatly thank all the Forage Breeding and Genetics Lab members and staff at the University of Florida Plant Science Research and Education Unit, Citra, FL for providing help during the field trial and data collection.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

## References

1. Quesenberry, K.; Munoz, P.; Blount, A.; Kenworthy, K.; Crow, W. Breeding forages in Florida for resistance to nematodes. *Crop Pasture Sci.* **2014**, *65*, 1192–1198. [CrossRef]
2. Jank, L.; Quesenberry, K.H.; Sollenberger, L.E.; Wofford, D.S.; Lyrene, P.M. Selection of morphological traits to improve forage characteristics of *Setaria sphacelata* grown in Florida. *N. Z. J. Agric. Res.* **2007**, *50*, 73–83. [CrossRef]
3. Acuña, C.A.; Blount, A.R.; Quesenberry, K.H.; Kenworthy, K.E.; Hanna, W.W. Tetraploid bahiagrass hybrids: Breeding technique, genetic variability and proportion of heterotic hybrids. *Euphytica* **2011**, *179*, 227–235. [CrossRef]
4. Wallau, M.O.; Sollenberger, L.E.; Vendramini, J.M.B.; Gomide, C.A.M.; Mullenix, M.K.; Quesenberry, K.H. Performance of limpograss breeding lines under various grazing management strategies. *Crop Sci.* **2016**, *56*, 3345–3353. [CrossRef]
5. Bouton, J.H. Breeding lucerne for persistence. *Crop Pasture Sci.* **2012**, *63*, 95–106. [CrossRef]

6. Michaud, R.; Lehman, W.F.; Rumbaugh, M.D. World distribution and historical development. *Alfalfa Alfalfa Improv.* **1988**, *29*, 25–91.
7. Annicchiarico, P. Alfalfa forage yield and leaf/stem ratio: Narrow-sense heritability, genetic correlation, and parent selection procedures. *Euphytica* **2015**, *205*, 409–420. [[CrossRef](#)]
8. North American Alfalfa Improvement Conference. Available online: <https://www.naaic.org/resource/importance.php> (accessed on 8 May 2020).
9. Crop Production Summary Report, USDA, NASS. 2018. Available online: [https://www.nass.usda.gov/Publications/Todays\\_Reports/reports/cropan19.pdf](https://www.nass.usda.gov/Publications/Todays_Reports/reports/cropan19.pdf) (accessed on 19 May 2020).
10. Belda, M.; Holtanová, E.; Halenka, T.; Kalvová, J. Climate classification revisited: From Köppen to Trewartha. *Clim. Res.* **2014**, *59*, 1–13. [[CrossRef](#)]
11. Peters, G.P.; Andrew, R.M.; Boden, T.; Canadell, J.G.; Ciais, P.; Le Quéré, C.; Marland, G.; Raupach, M.R.; Wilson, C. The challenge to keep global warming below 2C. *Nat. Clim. Chang.* **2013**, *3*, 4–6. [[CrossRef](#)]
12. Hanberry, B.B.; Fraser, J.S. Visualizing current and future climate boundaries of the conterminous United States: Implications for forests. *Forests* **2019**, *10*, 280. [[CrossRef](#)]
13. Blount, A.R. *Forage Production in the Southern Coastal Plain*; Agronomy Department, IFAS Extension, Univ. of Florida: Gainesville, FL, USA, 2017; pp. 1–2. Available online: <https://edis.ifas.ufl.edu/aa265> (accessed on 8 May 2020).
14. Horner, E.S. Registration of Florida 66 Alfalfa (Reg. No. 48). *Crop Sci.* **1970**, *10*, 456. [[CrossRef](#)]
15. Horner, E.S.; Ruelke, O.C. Registraton of Florida 77 Alfalfa (Reg. No. 99). *Crop Sci.* **1981**, *21*, 797. [[CrossRef](#)]
16. Vilela, D.; Basigalup, D.H.; Juntolli, F.V.; Ferreira, R.P. Research priorities and future of alfalfa in Latin America. In Proceedings of the Second World Alfalfa Congress, Cordoba, Argentina, 11–14 November 2018; pp. 140–143. Available online: <https://repositorio.inta.gob.ar/handle/20.500.12123/4031> (accessed on 8 May 2020).
17. de Assis, G.M.L.; Ruggieri, A.C.; Mercadante, M.E.Z.; De Camargo, G.M.F.; Carneiro Júnior, J.M. Selection of alfalfa cultivars adapted for tropical environments with repeated measures using PROC MIXED of SAS®System. *Plant Genet. Resour.* **2010**, *8*, 55–62. [[CrossRef](#)]
18. Annicchiarico, P.; Scotti, C.; Carelli, M.; Pecetti, L. Questions and avenues for lucerne improvement. *Czech J. Genet. Plant Breed.* **2010**, *46*, 1–13. [[CrossRef](#)]
19. Bingham, E.T.; Goose, R.W.; Woodfield, D.R.; Kidwell, K.K. Complementary gene interactions in alfalfa are greater in autotetraploids than diploids. *Crop Sci.* **1994**, *34*, 823–829. [[CrossRef](#)]
20. Brummer, E.C. Capturing heterosis in forage crop cultivar development. *Crop Sci.* **1999**, *39*, 943–954. [[CrossRef](#)]
21. Li, X.; Brummer, E.C. Applied genetics and genomics in alfalfa breeding. *Agronomy* **2012**, *2*, 40–61. [[CrossRef](#)]
22. Annicchiarico, P.; Barrett, B.; Brummer, E.C.; Julier, B.; Marshall, A.H. Achievements and challenges in improving temperate perennial forage legumes. *CRC Crit. Rev. Plant Sci.* **2015**, *34*, 327–380. [[CrossRef](#)]
23. Biazzi, E.; Nazzicari, N.; Pecetti, L.; Brummer, E.C.; Palmonari, A.; Tava, A.; Annicchiarico, P. Genome-wide association mapping and genomic selection for alfalfa (*Medicago sativa*) forage quality traits. *PLoS ONE* **2017**, *12*, e0169234. [[CrossRef](#)]
24. Li, X.; Wei, Y.; Acharya, A.; Hansen, J.L.; Crawford, J.L.; Viands, D.R.; Michaud, R.; Claessens, A.; Brummer, C.E. Genomic prediction of biomass yield in two selection cycles of a tetraploid alfalfa breeding population. *Plant Genome* **2015**, *8*, 1–10. [[CrossRef](#)]
25. Brummer, E.C.; Casler, M.D. Cool-Season Forages. *Yield Gains Major US Field Crop* **2014**, *1108*, 33–52.
26. Bowley, S.R.; Christie, B.R. Inheritance of Dry Matter yield in a Heterozygous Population of Alfalfa. *Can. J. Plant Sci.* **1981**, *61*, 313–318. [[CrossRef](#)]
27. Riday, H.; Brummer, E.C. Crop breeding, genetics & cytology: Heterosis in a broad range of alfalfa germplasm. *Crop Sci.* **2005**, *45*, 8–17.
28. Casler, M.D.; Brummer, E.C. Theoretical expected genetic gains for among-and-within-family selection methods in perennial forage crops. *Crop Sci.* **2008**, *48*, 890–902. [[CrossRef](#)]
29. Piepho, H.P.; Möhring, J.; Melchinger, A.E.; Büchse, A. BLUP for phenotypic selection in plant breeding and variety testing. *Euphytica* **2008**, *161*, 209–228. [[CrossRef](#)]
30. John, J.A.; Williams, E.R. The Construction of Efficient Two-Replicate Row-Column Designs for Use in Field. *Appl. Statist.* **1997**, *46*, 207–214. [[CrossRef](#)]

31. Gilmour, A.R. Average information residual maximum likelihood in practice. *J. Anim. Breed. Genet.* **2019**, *136*, 262–272. [CrossRef]
32. R Development Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Stat. Comput.: Vienna, Austria, 2018; Available online: <https://www.R-project.org/> (accessed on 8 May 2020).
33. Amadeu, R.R.; Cellon, C.; Olmstead, J.W.; Garcia, A.A.F.; Resende, M.F.R.; Muñoz, P.R. AGHmatrix: R package to construct relationship matrices for autotetraploid and diploid species: A blueberry example. *Plant Genome* **2016**, *9*, 1–10. [CrossRef]
34. Smith, B.A.; Stringer, J.K.; Wei, X.; Cullis, B.R. Varietal selection for perennial crops where data relate to multiple harvests from a series of field trials. *Euphytica* **2007**, *157*, 253–266. [CrossRef]
35. Andrade, V.T.; Gonçalves, F.M.; Nunes, J.A.R.; Botelho, C.E. Statistical modeling implications for coffee progenies selection. *Euphytica* **2016**, *207*, 177–189. [CrossRef]
36. Schwarz, G. Estimating The Dimension of a Model. *Ann. Stat.* **1978**, *6*, 461–464. [CrossRef]
37. de Mendiburu, F. *Agricolae: Statistical procedures for agricultural research. R Packag. Version* **2014**, *1*.
38. Wickham, H. *Ggplot2: Elegant Graphics for Data Analysis*; Springer: Berlin/Heidelberg, Germany, 2016; ISBN 3319242776.
39. FAO. *The State of the World's Land and Water Resources for Food and Agriculture*; Food and Agriculture Organization of the United Nations: London, UK, 2011; Available online: <http://www.fao.org/docrep/017/i1688e/i1688e.pdf> (accessed on 11 May 2020).
40. Ugo, P.C.; Baker, D.; Morgan, N.; Ly, C.; Nouala, S. Investing in African Livestock: Business opportunities in 2030–2050. 2013. Available online: <http://www.fao.org/3/al757e/al757e.pdf> (accessed on 11 May 2020).
41. Basigalup, D.H.; Rossanigo, R.O.; Rodríguez, N.E.; Spada, M.; Del, C.; Collino, D.J.; Dardanelli, J.L.; De Luca, M.J.; Racca, R.W.; González, N.S.; et al. *El Cultivo De La Alfalfa En La Argentina*; Ediciones INTA: Buenos Aires, Argentina, 2007; 476p, ISBN 9875212428.
42. Smith, S.R.; Bouton, J.H. Selection within Alfalfa Cultivars for Persistence under Continous Stocking. *Crop Sci.* **1993**, *33*, 1321–1328. [CrossRef]
43. Falconer, D.S.; Mackay, T.F.C. *Introduction to Quantitative Genetics*, 4th ed.; Longman: Essex, UK, 1996; p. 464.
44. Brummer, E.C.; Shah, M.M.; Luth, D. Reexamining the relationship between fall dormancy and winter hardiness in alfalfa. *Crop Sci.* **2000**, *40*, 971–977. [CrossRef]
45. Riday, H.; Brummer, E.C. Narrow Sense Heritability and Additive Genetic Correlations in Alfalfa subsp. falcata. *J. Iowa Acad. Sci. IJAS* **2007**, *114*, 28–34.
46. Rios, E.F.; Kenworthy, K.E.; Gezan, S.A.; Munoz, P.R. Genetic Parameters for Phenotypic Traits in Annual Ryegrass. *Crop Sci.* **2019**, *59*, 2128–2140. [CrossRef]
47. Anthony, D.; Elbaum, S.; Lorenz, A.; Detweiler, C. On crop height estimation with UAVs. In Proceedings of the 2014 IEEE/RSJ International Conference on Intelligent Robots and Systems, Chicago, IL, USA, 14–18 September 2014; 2014; pp. 4805–4812. [CrossRef]
48. Berhongaray, G.; Basanta, M.; Jauregui, J.M. Water table depth affects persistence and productivity of alfalfa in Central Argentina. *Field Crops Res.* **2019**, *235*, 54–58. [CrossRef]

