



Article

Phenotypic and Nodule Microbial Diversity among Crimson Clover (*Trifolium incarnatum* L.) Accessions

Virginia Moore ^{1,2,*} , Brian Davis ³ , Megan Poskaitis ⁴ , Jude E. Maul ² ,
Lisa Kissing Kucek ⁵ and Steven Mirsky ²

¹ Department of Crop and Soil Sciences, North Carolina State University (NCSU), Raleigh, NC 27695, USA

² Beltsville Agricultural Research Center (BARC), United States Department of Agriculture (USDA), Agricultural Research Service (ARS), Beltsville, MD 20705, USA; jude.maul@usda.gov (J.E.M.); steven.mirsky@usda.gov (S.M.)

³ Department of Environmental Science and Technology, University of Maryland (UMD), College Park, MD 20742, USA; brianwdavis@gmail.com

⁴ School of Plant and Environmental Sciences, Virginia Polytechnic Institute and State University (VT), Blacksburg, VA 24061, USA; megankposkaitis@vt.edu

⁵ Dairy Forage Research Center (DFRC), United States Department of Agriculture (USDA), Agricultural Research Service (ARS), Madison, WI 53706, USA; lisa.kucek@usda.gov

* Correspondence: vmmoore3@ncsu.edu; Tel.: +1-301-504-0069

Received: 13 August 2020; Accepted: 8 September 2020; Published: 21 September 2020



Abstract: Crimson clover (*Trifolium incarnatum* L.) is the most common legume cover crop in the United States. Previous research found limited genetic variation for crimson clover within the National Plant Germplasm System (NPGS) collection. The aim of this study was to assess the phenotypic and nodule microbial diversity within the NPGS crimson clover collection, focusing on traits important for cover crop performance. Experiments were conducted at the Beltsville Agricultural Research Center (Maryland, USA) across three growing seasons (2012–2013, 2013–2014, 2014–2015) to evaluate 37 crimson clover accessions for six phenotypic traits: fall emergence, winter survival, flowering time, biomass per plant, nitrogen (N) content in aboveground biomass, and proportion of plant N from biological nitrogen fixation (BNF). Accession effect was significant across all six traits. Fall emergence of plant introductions (PIs) ranged from 16.0% to 70.5%, winter survival ranged from 52.8% to 82.0%, and growing degree days (GDD) to 25% maturity ranged from 1470 GDD to 1910 GDD. Biomass per plant ranged from 1.52 to 6.51 g, N content ranged from 1.87% to 2.24%, and proportion of plant N from BNF ranged from 50.2% to 85.6%. Accessions showed particularly clear differences for fall emergence and flowering time, indicating greater diversity and potential for selection in cover crop breeding programs. Fall emergence and winter survival were positively correlated, and both were negatively correlated with biomass per plant and plant N from BNF. A few promising lines performed well across multiple key traits, and are of particular interest as parents in future breeding efforts, including PIs 369045, 418900, 561943, 561944, and 655006. In 2014–2015, accessions were also assessed for nodule microbiome diversity, and 11 genera were identified across the sampled nodules. There was large variation among accessions in terms of species diversity, but this diversity was not associated with observed plant traits, and the functional implications of nodule microbiome diversity remain unclear.

Keywords: cover crop; crimson clover; germplasm diversity; legume

1. Introduction

Cover crops provide a wide range of ecosystem services including nutrient retention, nitrogen fixation, erosion control, weed suppression, improved soil structure and ecology, improved water

quality, and beneficial insect conservation [1–7]. Despite these widely acknowledged benefits, adoption has remained limited [8–10]. This is in part due to limited improvements in cover crop genetics.

To increase farmer adoption, there is a need for germplasm improvement in cover crops. A lack of adapted cover crop cultivars (e.g., with enhanced biomass and winter hardiness for various climates) presents a significant barrier to cover crop adoption [11,12]. Historically, plant breeding has focused resources and effort on breeding for increased productivity in cash crops, with remarkable success [13–19]. By comparison, relatively little breeding has focused on ecosystem services (e.g., improving crops for use as cover crops) [20]. Shifting resources towards cover crop breeding could result in significant improvement in cover crop germplasm and ultimately in increased adoption on the landscape. Many species used as cover crops have been bred as cash crops (e.g., for forage production). However, to initiate breeding for use as cover crops, it is necessary to evaluate existing variation for traits of interest specific to cover crop production. Previous studies have assessed the genetic and phenotypic variation available through the National Plant Germplasm System (NPGS) for cover crop species, notably hairy vetch (*Vicia villosa* L.) [21,22].

Crimson clover (*Trifolium incarnatum* L.) is the most common legume cover crop in the United States [10]. Valued for its biomass production and nitrogen contribution, crimson clover is used across many regions; it is commonly used as a forage in the Southeast US and can be grown as a winter or summer annually depending on the climate [23]. A previous screening of the genetic variation in the NPGS crimson clover collection found limited genetic diversity and a substantial overlap in pedigrees [24]. Modern US cultivars show especially low genetic variation, with nearly all cultivars derived from ‘Dixie.’ However, an assessment of phenotypic variation in the NPGS crimson clover collection has not been conducted for cover crop traits of importance to farmers: nitrogen fixation, biomass production, winter hardiness, and flowering time [12]. Further, identifying which microbes colonize nodules may provide insights into variation in biological nitrogen fixation among crimson clover genotypes. In other legume species, nodule occupancy has been observed to differ both diversity and richness in rhizobia species [25,26]. To improve crimson clover germplasm for cover cropping, the phenotypic variations in existing collections needs to be assessed for traits of interest.

In this study, we assessed phenotypic variation within the NPGS crimson clover collection, and also explored the genetic diversity of nodule microbial populations regulating biological nitrogen fixation (BNF). We evaluated the NPGS crimson clover collection for emergence, winter survival, flowering time, biomass per plant, nitrogen (N) content, and proportion of plant N from BNF. We also evaluated the diversity of nodule microbial populations among accessions, with the understanding that ultimately BNF will be a result of interactions between traits of the resident soil/inoculum population and plant traits that affect nodule entry selectivity, nodule number, and activity.

2. Materials and Methods

2.1. Plant Material

Germplasm consisted of 37 accessions of crimson clover, representing all plant introduction (PI) lines of crimson clover in the NPGS with seed available for distribution. However, seven of these accessions were not available from NPGS after the 2012–2013 season, so the remaining 30 accessions were evaluated in 2013–2014 and 2014–2015 (Table S1).

2.2. Field Evaluation

Phenotypic evaluation of crimson clover accessions was conducted over three growing seasons, with crimson clover direct-seeded by hand in late September 2012, 2013, and 2014 in field plots in Beltsville, MD, USA (39°02′ N, 76°56′ W) on sandy loam soils (Russett Christiana complex). The field design was a randomized complete block design (RCBD) with four replications in each year, except for five accessions planted in 2015, which only had three replications due to limited seed availability.

Each plot was a single row of 0.6 m in length with 1.5 m between plots. Between 37 and 45 seeds were planted per plot, depending on seed availability in each year.

Fall emergence was evaluated in late October of each year by counting the total number of plants in each plot. Winter survival was determined by counting the total number of plants per plot in late April divided by the total number of plants counted in the fall. Flowering time was evaluated by recording percent flowering on a per-plot basis on a scale from 0% (no flower buds present) to 100% (all flowers dried up entire length of head). Flowering evaluations took place periodically between late April and early June. In 2013, evaluation took place on six dates: 23 April, 9 May, 15 May, 24 May, 30 May, and 4 June. In 2014, evaluation took place on five dates: 28 April, 6 May, 13 May, 19 May, and 27 May. In 2015, evaluation took place on eight dates: 25 April, 29 April, 4 May, 7 May, 11 May, 14 May, 18 May, and 21 May. The frequency of evaluations and total duration of evaluation period varied from year-to-year primarily due to the effects of year-to-year weather variation (Table 1) on the rate of growth and development.

Table 1. Total accumulated growing-degree days (GDD), freezing-degree days (FDD), days below freezing without snow cover, total precipitation (mm), and minimum low temperature (°C) in Beltsville, MD, from crimson clover planting to final data collection. Weather data were acquired from an on-site weather station at the Beltsville Agricultural Research Center (BARC).

Measurement	Growing Season		
	2012–2013	2013–2014	2014–2015
Growing degree days (GDD)	2208.9	2102.8	2198.2
Freezing degree days (FDD)	84	115	102
Days below freezing without snow cover	80	86	72
Total precipitation (mm)	596.0	802.8	616.2
Minimum low temperature (°C)	−11.0	−16.5	−17.0

Once an accession was rated at 50% or greater for flowering, biomass was collected. All plants in the plot were pulled up with roots attached. Plants were counted and the roots were clipped. All plants within a plot were placed in the same brown paper bag and dried. Dry weight was recorded and plants were ground for laboratory evaluation of nitrogen content, proportion of nitrogen from BNF, and metagenomic analysis.

2.3. Laboratory Evaluation

The crimson clover biomass samples were separated into shoots and roots. Shoots were oven-dried (60 °C) for approximately 72 h, weighed, and ground to pass a 1.0-mm screen. Tissue C and N concentrations and ¹⁵N natural abundance were determined for the shoot material of each accession using a Thermo Delta V Isotope Ratio Mass Spectrometer (Thermo Scientific, Waltham, MA, USA) and Carlo Erba NC2500 Elemental Analyzer (Carlo Erba, Milan, Italy). Isotopic abundance data were expressed as $\delta^{15}\text{N}$ in parts per thousand (‰), representing the abundance of plant tissue ¹⁵N relative to that of atmospheric N₂.

For 2015 samples, fresh roots were prepared for nodule collection by carefully removing soil from the root system, breaking as few roots as possible. Nodules were washed in a beaker of deionized water, placed onto a 53- μm sieve, and washed with deionized water until clean of soil and organic matter. The nodules were surface-sterilized by transferring to Falcon tubes containing 10% (v/v) bleach, gently inverting the tubes over a five-minute period, and washing again with deionized water until bleach odor was gone.

For each crimson clover accession, 500 mg of surface sterilized nodules were placed in a 1.5 mL Eppendorf tube and submerged in 1× phosphate buffered solution (PBS). Three to four sterilized nodules from each preparation were reserved before mashing and plated on yeast mannitol broth (YMB) to serve as controls for the surface sterilization procedure (no growth observed). The remaining nodules were mashed using an Eppendorf micropestle. The mashed nodules were centrifuged for

two minutes at 14,000 rpm to pellet the residual plant biomass. A total of 100 µL of mashed nodules were added to two mL of yeast mannitol broth YMB in sterile, 4.5 mL culture tubes and incubated in the dark on a shaker at 28 °C and 225 rpm for four d. Then, 100 µL of this culture was plated onto YMB plates and incubated at 28 °C for four d. After four days, a bacterial lawn developed on all plates. To collect the cells for analysis, 300 µL of 1× PBS buffer was added to the surface of each plate using a cell scraper, and the bacterial biomass from the surface of the plate was scraped and transferred into a single Eppendorf tube. The bacterial biomass was pelleted and divided into two aliquots; one was immediately prepared for storage as a glycerol stock. The other half was frozen at −20 °C until it was used for genomic DNA extraction and downstream molecular analysis.

The nodule rhizobacteria from 19 accessions were individually processed for genomic DNA extraction using Qiagen power plant genomic DNA (Qiagen Inc., Germantown, MD, USA). Each metagenomic extraction was individually barcoded using the Illumina 30 barcodes of a 96-index bar code kit following manufacturer's instructions (Illumina, Inc., San Diego, CA, USA). The indexed metagenomic libraries were then sequenced on an Illumina nextSeq 500 following manufacturer's instructions using all channels of a flow cell.

2.4. Data Analysis

Six phenotypic performance metrics were assessed: fall emergence, winter survival, flowering time, biomass per plant, nitrogen content, and proportion of plant nitrogen from BNF. For each response variable, a linear mixed model was fit, with accession, year, and year × accession as fixed effects and replicate within year as random effects. Where appropriate, log-transformation was applied to the response variable to achieve homogeneity of variances. For each model, means and least-significant difference (LSD) estimates were extracted.

Fall emergence was calculated as the count of living plants as a proportion of seeds planted. Winter survival was calculated as the count of living plants in the spring as a percent of fall-emerged plants. When more plants were present in the spring than in the fall (e.g., due to spring emergence), winter survival was set to 100%.

Flowering time was evaluated as the growing-degree days (GDD) required to reach 25% flowering on a per-plot basis. GDD was calculated using a base temperature of 0 °C [27]. For plots that reached 25% flowering before the first rating (17 of 323 plots), GDD to 25% was set to the GDD of the first observation date. Likewise, for plots that never reached 25% flowering (5 of 323 plots), GDD to 25% was set to that of the last observation date. For plots that reached 25% flowering between observation dates, a linear interpolation function was used to estimate the GDD at which 25% flowering occurred.

Biomass for each plot was calculated on a per-plant basis by dividing the total plot biomass by the number of plants present in the plot at biomass harvest. The biomass and plant-tissue %N content data showed a non-normal distribution, so a log-transformation was applied. The proportion of tissue-N derived from BNF was estimated using the ¹⁵N natural abundance method [28] and statistically analyzed as above. All analyses were conducted in R version 3.6.1 (R Foundation for Statistical Computing, Vienna, Austria) and used package {nlme} (3.1-143) for model fitting [29]. Scripts for data analysis and raw data for the plant phenotypic analysis are available at <https://data.nal.usda.gov/dataset/data-phenotypic-and-nodule-microbial-diversity-among-crimson-clover-trifolium-incarnatum-l-accessions>.

All metagenomic analyses were conducted with the free on-line bioinformatics suite Kbase (<https://kbase.us/>). Paired-end reads were uploaded to the Kbase cloud then joined using the join-paired-ends import tool, adapters and indices were trimmed, and joined reads with a Phred quality score over 30 were kept for further analysis and processed separately from this point forward. Paired end libraries were assembled using KBaseGenomeAnnotations.Assembly-5.0, which employs Megahit 1.1 [30] with customized parameter setting. Parameters were set to min-count multiplicity for filtering (k_min+1)-mers, default 2, k-min kmer size (≤127), default 21, k-max kmer size (≤255), default 141, k-step increment of kmer size of each iteration (≤28), default 12, k-list list of kmer size (range 15–255, increment ≤28), min-contig-len minimum length of contigs to output (≥300), default 300.

Assembled contigs were then annotated using Rapid Annotations and the Subsystems Technology toolkit (RASTtk) [31–33]. Default search and alignment parameters are defined at <http://rast.nmpdr.org/>. Annotated metagenomic libraries were compared to each other using GenomeComparisonSDK v.0.0.4 (<https://kbase.us/>), which uses a k-mer approach to compute a protein pangenome. This approach builds on the sequence-based homologous protein families identified in the input Pangenome object by adding to the accumulated protein family annotation by identifying putatively similar functions within or among family groups. In addition, we identified the genes for tRNAs, rRNAs, and transposons. This database was then used to identify all the co-occurring protein families, tRNAs, rRNAs, and transposons among all nodule metagenomes. Scripts for data analysis and raw data for the metagenomic analysis are available at <https://narrative.kbase.us/narrative/26966>.

3. Results

3.1. Fall Emergence

The effects of accession, year, and accession \times year were all significant to at least the 0.05 level. Replicate explained a small amount of variance in the model compared with accession, year, and accession \times year interaction (Table 2). Fall emergence averaged across years varied from 16.0% (PI 233812) to 70.5% (PI 556990) of seeds planted (Figure 1), with a grand mean of 48.6% of plants emerged across all genotypes. Emergence differences were likely caused by a combination of genetic and seed source effects on seed coat impermeability and seed viability [34,35]. The effect of year on emergence was inconsistent across genotypes (indicated by the significant interaction effect), but there were several genotypes with consistently high or low emergence. The following nine accessions had a mean fall emergence of 60% or greater, and merit consideration as breeding parents: PIs 369045, 556990, 561942, 561943, 561944, 613042, 613043, 613047, and 655006. If low emergence was due to seed coat impermeability rather than non-viable seed, accessions with low fall emergence may also be valuable in a breeding context, as some producers prefer to use hard-seeded crimson clover as a self-seeding forage in pastures. Conversely, other producers prefer lines with low hard-seed for consistent emergence as a cover crop in rotations with annual cash crops [24,34]. Further investigation is needed to determine the effect of genetic and environmental variance on hard seed and seed viability: accessions that exhibited high and low emergence in this study (Figure 1) could be grown in diverging environments, then harvested seed could be assayed for impermeability. As previous research has found hard seed to be heritable [36], divergent selection for hard seed could be a promising breeding target for crimson clover.

Table 2. Results of the analysis of variance for each of the six phenotypic performance metrics: fall emergence, winter survival, flowering time, biomass per plant, nitrogen content, and plant nitrogen from BNF.

Term	Fall Emergence	Winter Survival	Flowering Time	Biomass per Plant	N Content	Plant N from BNF
F value						
Accession	49.30 **	30.63 ***	2819.88 ***	22.66 ***	140.30 ***	21.61 ***
Year	4.30 *	144.41 ***	16.58 ***	317.17 ***	92.43 ***	77.34 ***
Accession \times Year	1.72 **	1.04	1.81 **	1.18	2.16 ***	2.32 ***
chi square						
Replicate (Year)	0.20	384.86 ***	1.10	357.61 ***	75.77 ***	443.61 ***
Variance decomposition						
Fixed effects	0.75	0.33	0.83	0.38	0.43	0.45
Random effects	0.02	0.72	0.03	0.74	0.37	0.87
Coefficient of Variation (CV)	0.20	0.23	0.03	0.34	0.13	0.08

p value significance is denoted at the following levels: * <0.05, ** <0.01, *** <0.001.

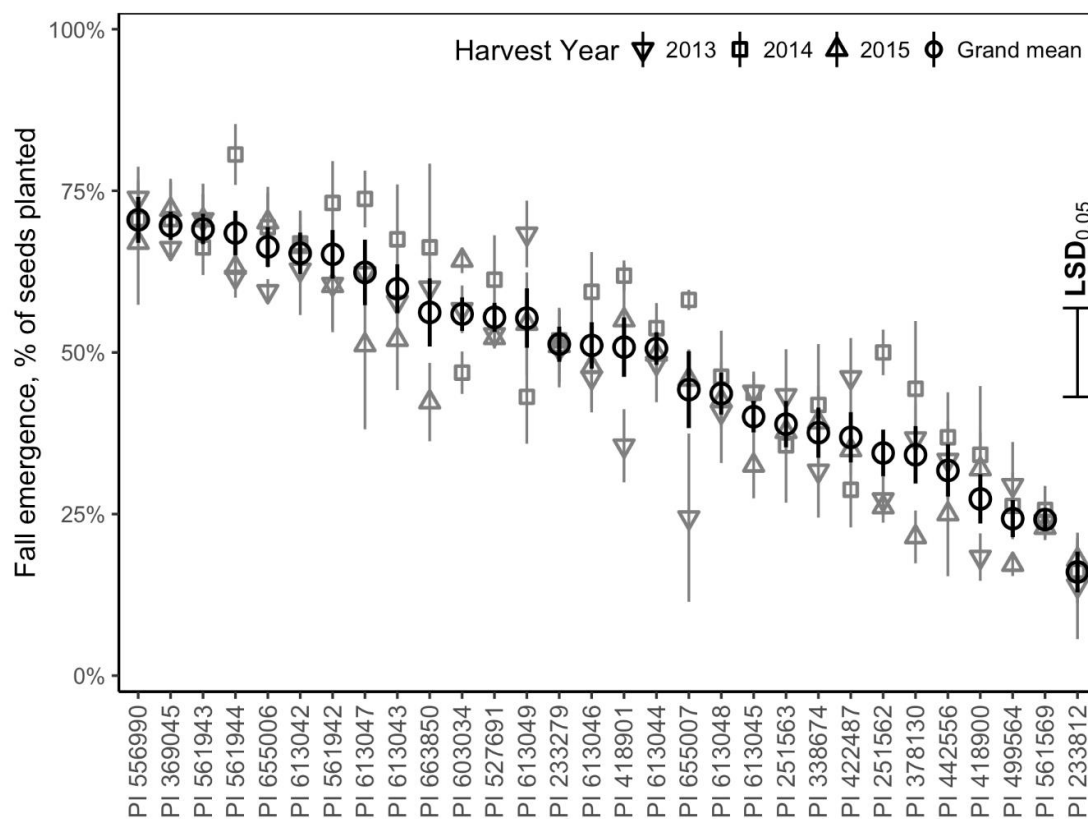


Figure 1. Effect of crimson clover accession on fall emergence in Beltsville, MD, evaluated as emerged plants per seeds planted. Grand mean, harvest year means, and least significant difference (LSD) estimates are displayed.

3.2. Winter Survival

The effects of accession, year, and replicate on winter survival were significant at the 0.001 level. Replicate explained more variance in the model than accession, year, and accession \times year interaction combined (Table 2). Winter survival averaged across years varied from 52.8% (PI 418900) to 82.0% (PI 613046) of seeds planted (Figure 2), and a grand mean of 69.6% of emerged plants survived the winter. The 2013–2014 season had lower winter survival than other years (few accessions surpassed a mean of 50% winter survival) but also much greater variability in survival. The reduced winter survival in 2013–2014 is likely explained by the greater number of FDD and days below freezing without snow cover (Table 1). The large contribution of replicate, relative to fixed effects, highlights challenges associated with selection for winter hardiness, especially the complex genetics and physiology of winter hardiness and its high degree of variability even within a single field and growing season [37,38]. However, the significant accession effect and the lack of significant accession \times year interaction suggests that agronomically important differences in winter hardiness in crimson clover may be observed even in years with relatively limited winter kill, and that selection can take place even in milder winters.

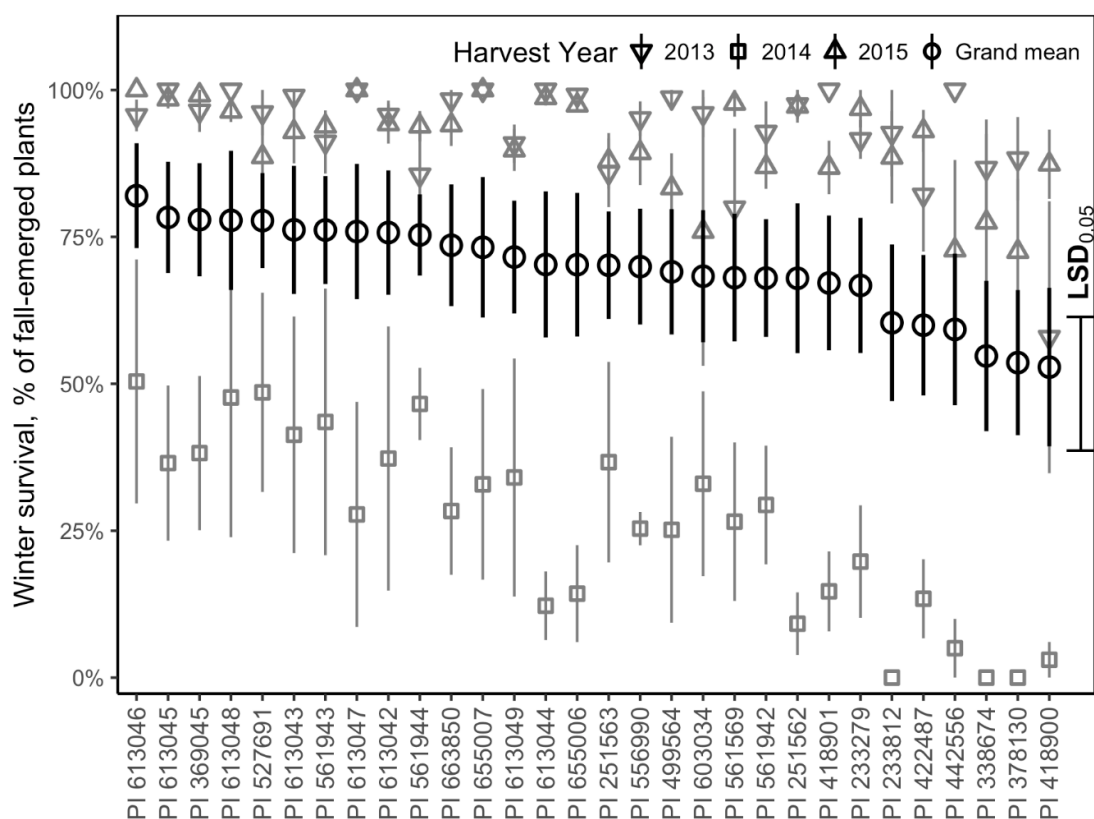


Figure 2. Effect of crimson clover accession on winter survival in Beltsville, MD, evaluated as surviving plants in the spring per plants emerged in the fall. Grand mean, harvest year means, and least significant difference (LSD) estimates are displayed.

3.3. Flowering Time

Effect of accession, year, and accession \times year interaction on crimson clover flowering time were significant to at least the 0.01 level. As with fall emergence, the replicate explained a small amount of variance in the model compared with accession, year, and accession \times year interaction (Table 2). GDD to reach 25% maturity varied from 1470 (PIs 561943 and 561944) to 1910 (PI 418901) (Figure 3), and accessions reached 25% maturity at a grand mean of 1581 GDD. Although the accession \times year interaction was significant, flowering time was relatively consistent across years except for the accessions with the most extreme values. This consistent performance aligns with previous studies in other clover species, which found high heritability for flowering time [39,40]. While the majority of accessions fell within a narrow band of flowering time (around 1600 GDD), three accessions (PIs 561942, 561943, and 561944) originating in the southeast US matured somewhat earlier, with all three reaching 25% maturity prior to 1500 GDD. These accessions represent potential parents for development of an early-maturing population. One accession, PI 418901, was significantly later-maturing than all other accessions; it reached 25% maturity at 1910 GDD, and the next-latest accession (PI 418900) reached 25% at 1690 GDD. This accession, therefore, is an excellent candidate for development of a late-maturing population. Both early- and late-maturing crimson clover may be of interest to farmers, depending on region and cropping system.

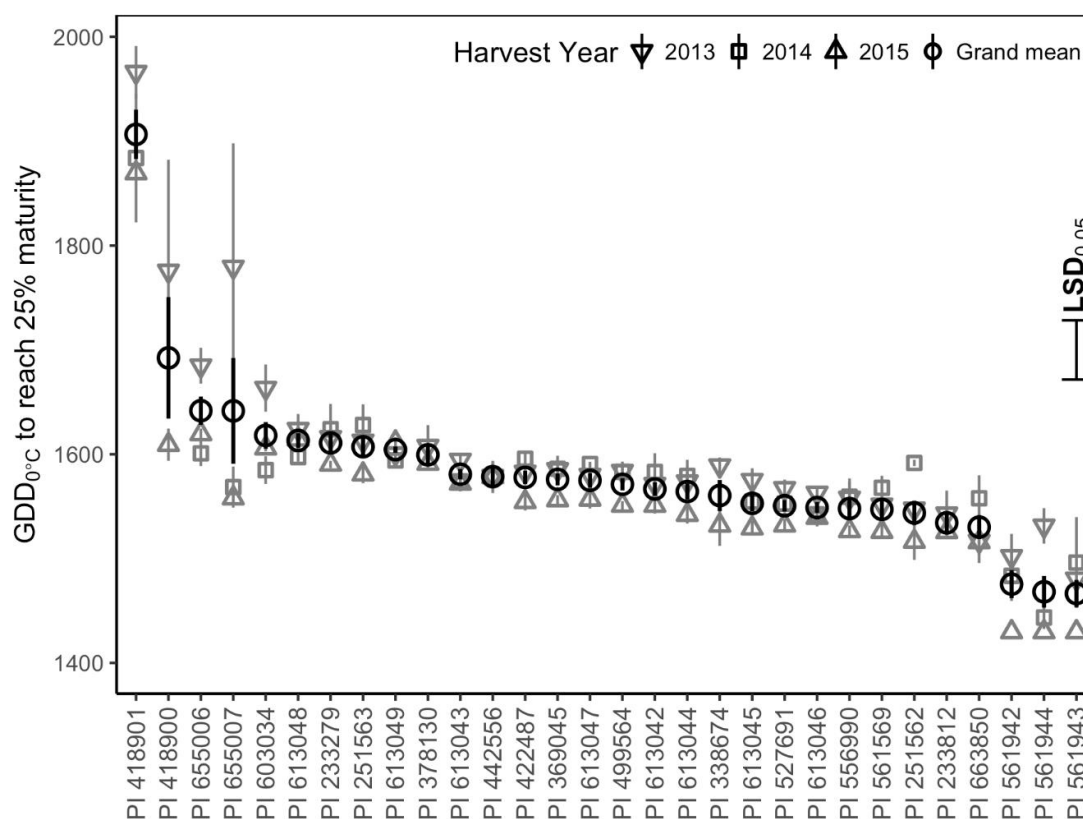


Figure 3. Effect of crimson clover accession on flowering time in Beltsville, MD, evaluated as growing degree days (GDD) to 25% flowering, calculated at a base temperature of 0 °C. Grand mean, harvest year means, and least significant difference (LSD) estimates are displayed.

3.4. Biomass per Plant

Effects of accession, year, and replicate on biomass per plant were significant to at least the 0.01 level. Replicate explained more variance in the model than accession, year, and accession \times year interaction combined (Table 2). Biomass per plant varied from 1.52 g plant⁻¹ (PI 233279) to 6.51 g plant⁻¹ (PI 442556) (Figure 4), and accessions produced a grand mean of 2.97 g plant⁻¹. Across accessions, plants produced less biomass in 2013–2014 than the other years, likely due to the colder winter and/or wetter year (Table 1). Again, the large contribution of replicates within the year highlights the impact of field spatial variability on plant biomass and underscores the challenge of phenotyping and selecting for increased biomass production. While some crimson clover cultivars have been developed with improved plant vigor and forage yields [34], given the limited variance explained by accession within the crimson clover collection, progress in biomass production is likely to be slow in a crimson clover cover crop breeding program as well. However, as for winter survival, accession \times year interaction is not significant, suggesting that agronomically important differences in biomass in crimson clover may be observed even in years with low biomass per plant overall.

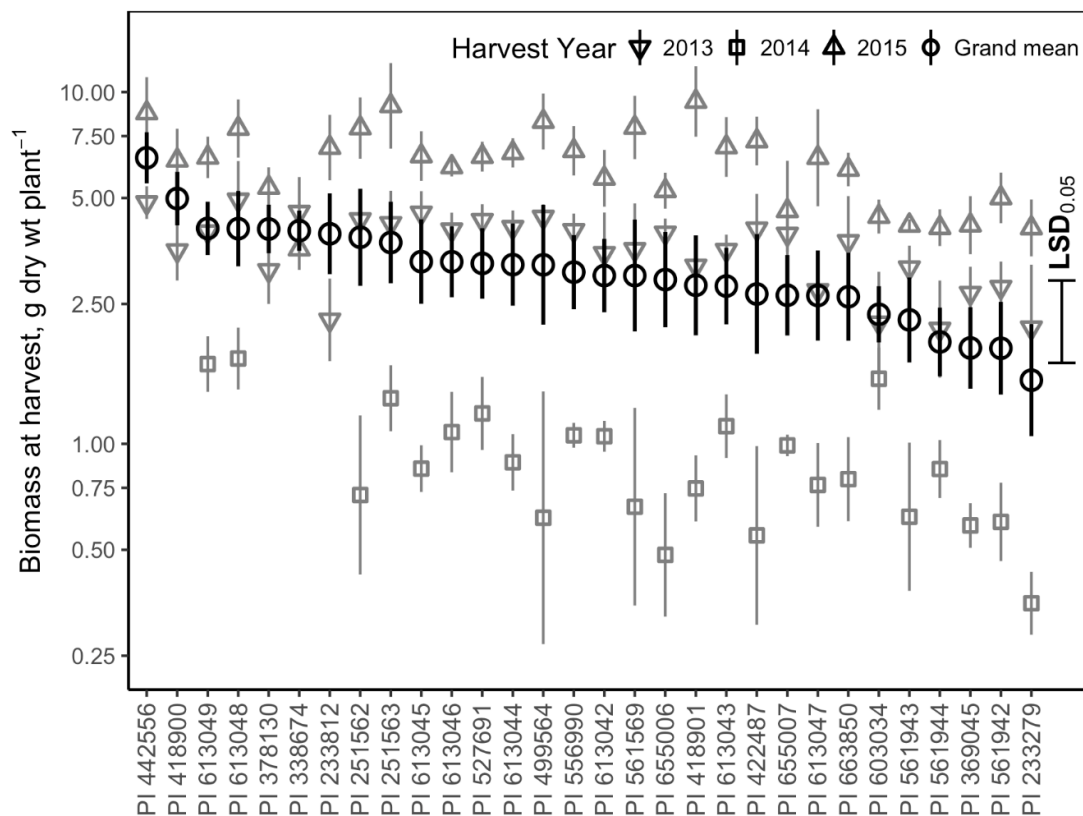


Figure 4. Effect of crimson clover accession on biomass per plant in Beltsville, MD, evaluated as dry weight per plant at harvest. Data are presented on a log scale. Grand mean, harvest year means, and least significant difference (LSD) estimates are displayed.

3.5. Nitrogen Content

The effects of accession, year, accession \times year interaction, and replicate on nitrogen content were all significant at the 0.001 level. Replicates explained comparable variance as accession, year, and accession \times year interaction combined (Table 2). Plant nitrogen content varied from 1.87% (PI 418901) to 2.24% (PI 233812) (Figure 5), and accessions contained a grand mean of 2.07% nitrogen. In 2013–2014, in which most accessions produced less biomass, many accessions also contained higher nitrogen content. The relationship between yield and crude protein content within cultivars due to seasonal variation has not been sufficiently addressed in the literature and merits further research. However, the range of nitrogen content present in the accessions is relatively narrow, and so in terms of total N contributed by the cover crop, an accession producing more biomass will likely produce more N even if its percent N is on the lower end of the observed spectrum. As with winter survival and biomass, replicate contributed a large portion of the model variance, again indicating that nitrogen content is highly impacted by spatial variability within the field.

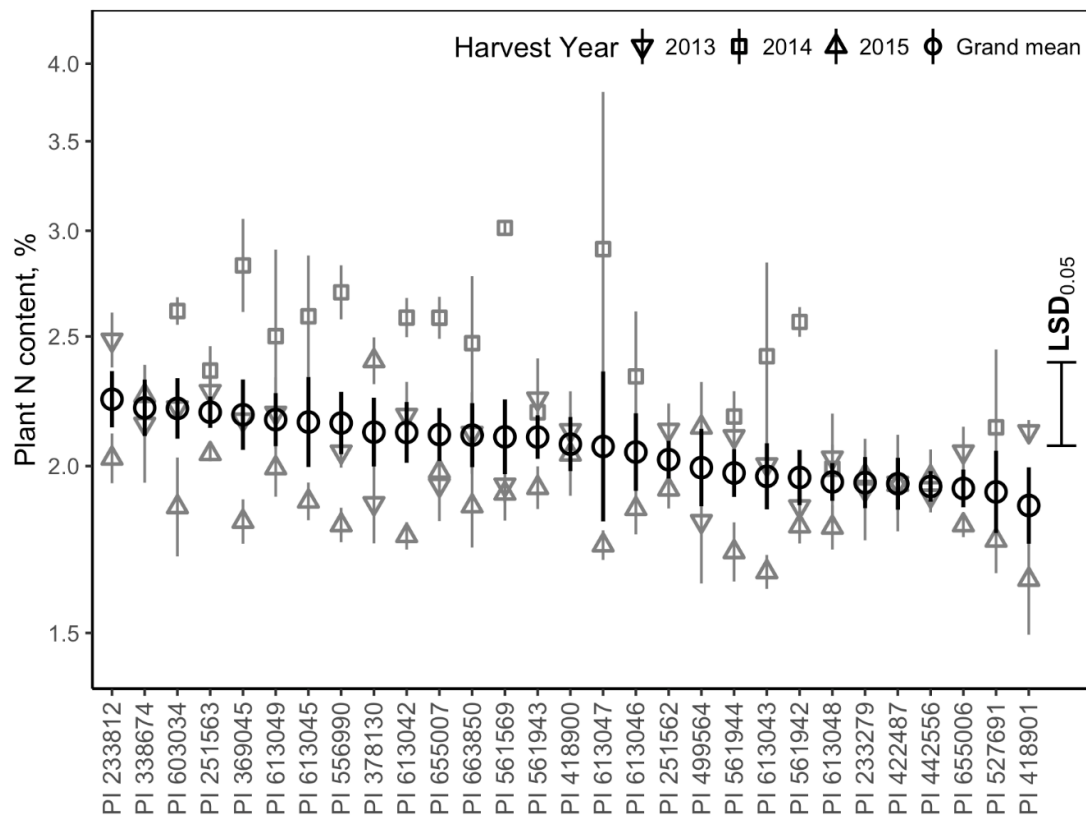


Figure 5. Effect of crimson clover accession on plant nitrogen content in Beltsville, MD, evaluated as percent N of dry weight. Data were log-transformed, but reverse-transformed for display purposes. Grand mean, harvest year means, and least significant difference (LSD) estimates are displayed.

3.6. Proportion of Plant Nitrogen as BNF

Again, the effects of accession, year, accession \times year interaction, and replicate on BNF were all significant at the 0.001 level. Again, the replicate explained more variance in the model than accession, year, and accession \times year interaction combined (Table 2). The proportion of plant nitrogen as BNF varied from 50.2% (PI 527691) to 85.6% (PI 655006) (Figure 6), and accessions contained a grand mean of 73.1% BNF as a proportion of plant nitrogen. However, the accessions with the highest mean proportion of N as BNF were accessions not included in 2013–2014, a year with lower BNF across all included accessions. Few breeding efforts have focused on nitrogen fixation ability in part due to challenges in implementing efficient screening systems, although differences between clover cultivars have been noted [41]. Again, although significant differences between accessions were observed, the replicate explained a larger portion of the variance than fixed effects, indicating that biological nitrogen fixation is highly impacted by field spatial variability and selection for this trait in a crimson clover cover crop breeding program is likely to be slow.

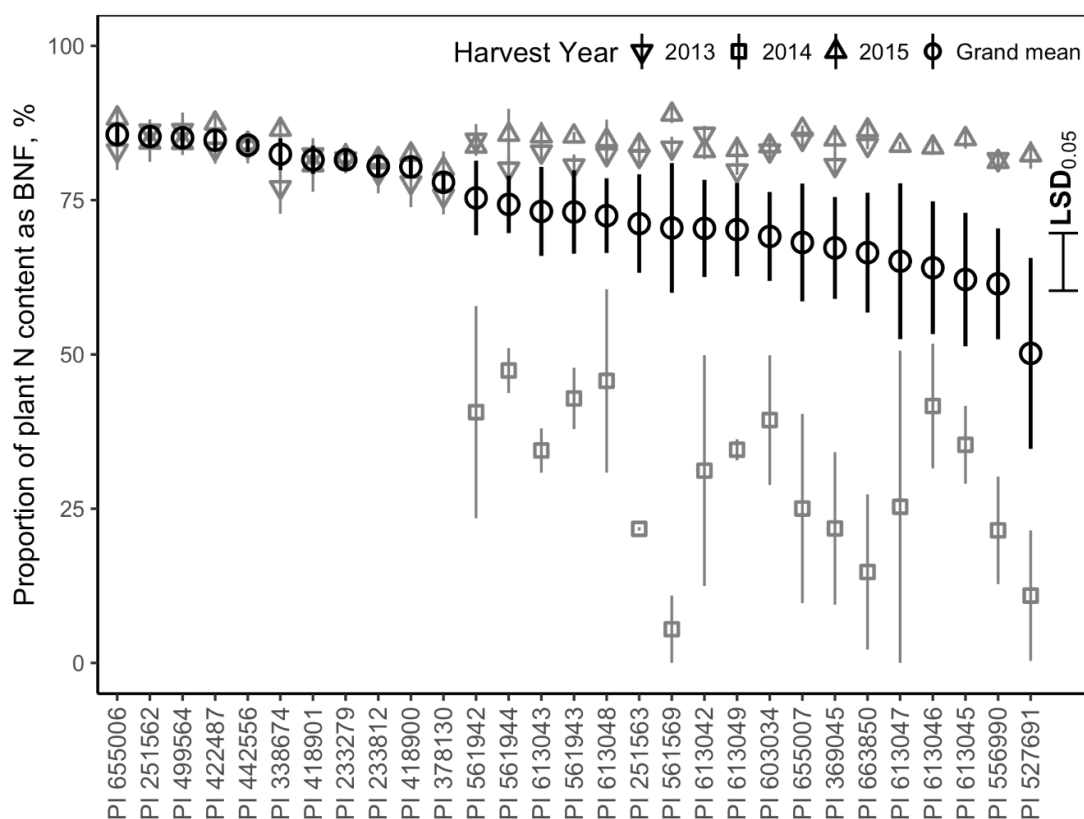


Figure 6. Effect of crimson clover accession on proportion of plant nitrogen content as biologically fixed nitrogen (BNF) in Beltsville, MD, evaluated as percent BNF per total N content. Grand mean, harvest year means, and least significant difference (LSD) estimates are displayed.

3.7. Nodule Metagenome

In total, we constructed nodule metagenomic libraries from 19 of the crimson clover accessions. Each library contained on average 8–10 million reads, equivalent to about 1.51×10^9 bases with an average single read length of 147 bp. Sequence assembly and duplicate sequence removal resulted in metagenomic libraries that ranged in size from 4.6 to 23.7 Mb. The number of contigs generated for each accession ranged in size among the libraries from 300 bp to 190 kb. The total number of contigs ranged from 5693 to 18,765, with a mean of 13,090.

The guanine-cytosine (GC) content of the libraries ranged from 50.4 to 62.8 % GC, which is, on average, lower than the three reference genomes that were used as comparisons (60.6, 60.9, and 61.1 %GC for *Rhizobium leguminosarum* Rt24.2, *Rhizobium leguminosarum* WSM1689, *Rhizobium leguminosarum* WSM2304, respectively) (Figure S1).

Homology to the rRNA operon was used to determine the diversity of bacterial species within each nodule metagenomic library at >99% identity. Other parts of the genome were identified to the genera or species level by alignment with documented gene fragments in KBASE with a minimum length > 1 kb, and an insertion or deletion gap size of 25 bp. All species identified had a minimum E-value of <0.001. These stringent conditions enabled the identification of many taxa at species level or variants contained within species. Among the nodule metagenomic libraries, the 11 genera identified were *Rhizobium*, *Mesorhizobium*, *Xanthomonas*, *Stenotrophomonas*, *Aquiflexum*, *Sinorhizobium*, *Bradyrhizobium*, *Agrobacterium*, *Paenibacillus*, *Sphingomonas*, and *Pseudomonas*. *Bacillus* dominated the number of reads that could be identified to a species level with reads specific enough to identify nine distinct *Bacillus* species in contigs of genomic regions greater than 10 kilobases. *Rhizobium* was clearly represented but is known to be more challenging to grow on media and therefore may have had a disadvantage on media relative to its importance in the nodule.

We identified 160,000 protein families, tRNA, ORFs, rRNA, and transposon genes; of these, 42% of the protein families were of unknown function. In the 11 bacterial genera identified in large reconstructed genomic contigs, all contained whole or partial complements of the nitrogen fixation complexes (nitrogenase genes, *nif*, and *nod*) as well as nitrogen responsive gene operons such as the *nor* and *nos* genes. Searches within the nodule metagenomes for other potential plant growth promoting rhizobacterial genes yielded numerous antibiotic genes as well as hydrogenases, which have been associated with increases in efficiency in biological nitrogen fixation.

We used three reference *Rhizobium leguminosarum* strains in our analysis (WSM1689, WSM2304, and Rt24.2 which have 6760, 6700, and 7080 annotated genes, respectively). The reference genomes have about 9500 identified elements that are over 1 kb in size and can be assigned to a specific species. The crimson clover accessions harbored a wide range of diversity within their nodule phytobiomes, from PI 369045, which contained 20,729 orthologous metagenome regions, to PI 561944, which contained 72,930 orthologous metagenome regions. The clover nodule microbiome with the lowest diversity of identifiable genetic regions (PI 369045) has about twice as much genetic diversity as the three reference genomes (Figure S1). The diversity of taxa identified among the nodules range between 12 and 20 different species of bacteria among the clover accessions. This diversity and the sequencing depth of our study (average library size) suggests that, for some accessions, we may not have data covering the full genomes of all genera within the nodule.

Several genes with roles in nitrogen fixation were found across multiple taxa identified within the crimson clover nodules (Table 3). All of the major taxa identified contained the *nifH* gene, which is a primary gene encoding the nitrogenase enzyme that carries out nitrogen fixation [42]. The *narG* nitrate reductase gene was found in two of the identified taxa, and has been shown to recycle fixed nitrogen through the dissimilatory nitrate reduction to ammonium (DNRA) pathway [43]. Anaerobic ammonium oxidation has been shown to remove significant amounts of nitrite and ammonium from ecosystems fixing nitrogen [44]. The *hlyD* gene, which is a multi-drug efflux transporter associated with antimicrobial resistance and survival in toxic environment, was found in three of the identified taxa. These types of excretion systems are likely to be involved with infection and nodule morphology [45]. The *hesB* gene, which is part of the nitrogenase, nitrogen fixation operon, was also found in three of the identified taxa [46]. Interestingly, the majority of research related to the role of *hesB* in BNF efficiency and regulation has occurred in algal systems and platonic freshwater cyanobacteria systems [47–49]. The possible roles of these genes in terrestrial leguminous systems has largely been ignored. There is potential to use metagenomic analysis of nodules to understand the efficiency of nitrogen fixation, recycling, and survival of symbionts in the nodule merit.

Table 3. Taxa identified in crimson clover nodules containing genes of interest related to nitrogen fixation. Genes of interest include *nifH* nitrogenase, *narG*- respiratory nitrate reductase, *hlyD*- multi drug efflux transporter, and *hesB*-, a regulatory gene that is only expressed during nitrogen fixation conditions.

Taxa Identified by 16S Operon	<i>nifH</i>	<i>narG</i>	<i>hlyD</i>	<i>hesB</i>
<i>Acidovorax</i>	X			
<i>Bacillus/Psychrobacillus/Paenibacillus</i>	X	X	X	X
<i>Microbacterium</i>	X			
<i>Pelamonas puraquae</i>	X			
<i>Propionibacterium</i>	X			X
<i>Pseudomonas (fluorescens, koreensis)</i>	X		X	
<i>Rhizobium</i>	X			
<i>Rhizobium leguminosarum</i>	X		X	X
<i>Sphingomonas</i>	X			
<i>Sporosarcina</i>	X	X		

3.8. Correlation among Traits

Associations among plant traits and nodule metagenome features were assessed using Pearson correlation. Among the plant phenotypic traits measured, only fall emergence and winter survival were positively correlated. Negative correlations were observed between fall emergence and biomass per plant, winter survival and biomass per plant, fall emergence and plant N from BNF, and winter survival and plant N from BNF. No correlations with any other traits were observed for flowering time or N content (Table 4).

Table 4. Correlation among plant phenotypic traits and nodule metagenomic library features.

	Fall Emergence	Winter Survival	Biomass per Plant	Flowering Time	N Content	Plant N from BNF	Library Size	Contigs	GC Content
Winter Survival	0.59 ***	1							
Biomass per Plant	−0.60 ***	−0.48 **	1						
Flowering Time	−0.15	−0.23	0.15	1					
N Content	−0.12	−0.09	0.06	−0.25	1				
Plant N from BNF	−0.42 *	−0.65 ***	0.24	0.24	−0.27	1			
Library Size	0.12	0.09	−0.09	0.05	0.07	0.18	1		
Contigs	−0.12	−0.05	−0.04	−0.35	−0.26	−0.06	−0.16	1	
GC Content	−0.10	0.12	−0.04	0.05	−0.26	0.29	0.46 *	−0.03	1
Microbial Diversity	−0.14	−0.20	0.17	−0.05	−0.10	0.25	0.60 **	0.37	0.58 **

p value significance is denoted at the following levels: * <0.05, ** <0.01, *** < 0.001.

Several accessions ranked in the top five accessions for two or three key traits. Of particular interest, PI 369045 was among the top accessions for emergence, survival, and N content. PI 418900 ranked second for both biomass and late flowering. PIs 561943 and 561944 were the earliest-flowering accessions and were both also among the top accessions for emergence. PI 613048 had both high winter survival and biomass per plant, and PI 655006 was among the top accessions for emergence, proportion of N as BNF, and late flowering (Figures 1–6).

Several positive correlations were observed among the nodule metagenome features. However, there were no correlations between any plant phenotypes and any nodule metagenome features.

4. Discussion

Plant breeding is a critical tool to increase cover crop use and efficacy on agricultural landscapes [11,12]. Crimson clover is the most widely used legume cover crop in the United States [10], and there is a need to improve the crimson clover germplasm for important cover cropping traits such as increased biomass and winter hardiness. Modern US cultivars have significant overlap in pedigrees, as most are derived from ‘Dixie,’ and limited genetic diversity found in a previous study of the NPGS crimson clover collection [24], but this was the first phenotypic assessment of the crimson clover collection for cover crop traits important to farmers.

Despite the limited genetic diversity previously observed, we did find some meaningful phenotypic variation for traits of agronomic importance for cover crop usage. Flowering time showed a large accession effect, and agronomically relevant diversity was observed among accessions. Its lack of correlation with other traits indicates that selection for flowering time may be possible without negative impacts on other agronomically important traits (e.g., due to linkage drag). Both early- and late-maturing populations have utility in different cropping systems. These results suggest that flowering time would be a relatively straightforward trait for selection in a crimson clover breeding program.

Fall emergence also showed a large accession effect, but further work is needed to determine whether this variation is caused by genetics or seed source effects. Low emergence may have been due to hard seededness, in which case accessions with both low and high emergence will be valuable for crimson clover breeding programs, depending on cropping system context [24,34].

Replicate effects explained the most variance in winter survival, biomass per plant, and proportion of plant N as BNF, highlighting the high degree of spatial variability among these traits and the challenges of phenotyping and selection [37,38,41,50,51]. However, unlike other traits, winter survival and biomass per plant did not exhibit accession \times year interaction, indicating that despite large annual variation in these traits, agronomically relevant differences in winter hardiness and biomass per plant are observable among crimson clover accessions even in years with mild winters or low biomass overall.

Additional challenges come with the negative correlations among key traits—both fall emergence and winter survival are negatively correlated with both biomass per plant and proportion of plant N as BNF. The presence of these correlations may limit breeders' ability to rapidly integrate all key agronomic traits in a single population. However, several accessions performed among the top five for multiple traits, and should be considered as parents for future breeding populations. All of the nodules contained more genetic diversity than would be expected if a single isolate occupied the nodules; the reference isolates had, on average, 6900 genetic elements [52], whereas the accessions surveyed ranged from 20,000 to 72,000 common genetic elements. This indicates that, in all of the accessions tested, we observed, from 3 to 10 times, the expected diversity of a single bacterial genome (reference genomes), but the range among accessions was quite broad. In the case of PI561944 (~72,000 elements) the genetic diversity is what might be observed if analyzing the genetic diversity of 9–12 genera of bacteria.

Additional analysis of the metagenome data showed that a large proportion of the identified bacterial taxa contained intact or partial nitrogenase operons [42]. There is evidence that genomes of nitrogen-fixing bacteria may have higher GC content than their non-fixing relatives [53]; the metagenomic GC content of the nodule microbiome may be an indicator of nodule populations that harbor a higher or lower proportion of nitrogen-fixing bacteria. Although we did not observe a significant relationship between GC content and BNF, our assessment was not quantitative in regards to extraction of total microbial biomass from the nodules.

Our data suggest that crimson clover accessions contain a wide range of bacteria in their nodules, but transferring these microbes from the nodule to a petri plate likely gave an advantage to the r-selected lineages of bacteria (i.e., those with high growth rates) that had successfully infected the nodule [54]. In addition, nodule metagenomic analysis was conducted in only one of the three growing seasons, so we are unable to draw conclusions about the stability of the nodule microbial populations or their impact on annual variation in agronomic traits. We also cannot yet clearly link the genetic diversity of nodule microbiomes of field grown clover to traits important in cover crop production such as biological nitrogen fixation and biomass production.

Although we cannot relate the metagenomic survey to the diversity of phenotypes analyzed for the field work portion of this study, this nodule microbiome dataset gives us a snapshot of the surveyed diversity of plant bacterial symbiosis in planta. Recent studies have shown that the nodule is much more diverse than previously thought; effectively managing the nodule microbiome will require an understanding of interactions among the diversity within the nodule in the context of the plant genotype [55–58]. It is not known whether this observed range of nodule metagenomic diversity is typical among closely related plant varieties, and additional research is needed to determine any salient relationships between microbial diversity and agronomic performance. The degree of annual and spatial variation in N content and BNF especially highlight the need to study environmental and management effects and genotype \times environment interactions on crimson clover microbial populations.

5. Conclusions

While the effect of crimson cover accession was significant across all evaluated traits, accessions showed particularly clear differences for fall emergence and flowering time, indicating greater diversity and potential for selection in cover crop breeding programs. By contrast, winter survival, biomass per plant, N content, and plant N from BNF showed greater within- and/or among-year variation. Fall emergence and winter survival were found to be positively correlated, and both of these traits

were negatively correlated with both biomass per plant and plant N from BNF. Several accessions were identified as high-performing across several key traits, and are of particular interest as parents in future breeding efforts: PIs 369045, 418900, 561943, 561944, and 655006. We observed larger than expected variation among accessions in terms of the diversity within the nodule microbial population, but did not observe any association between nodule metagenome features and plant phenotypic traits. Therefore, additional research is needed to determine functional implications of nodule microbiome diversity, and thus whether to incorporate such data into crimson clover breeding programs.

Supplementary Materials: The following are available online at <http://www.mdpi.com/2073-4395/10/9/1434/s1>. Figure S1: Metagenomic analysis of nodule endophytes; Table S1: Crimson clover accessions.

Author Contributions: Conceptualization, S.M. and J.E.M.; methodology, S.M., L.K.K., and J.E.M.; formal analysis, B.D., V.M., and J.E.M.; data curation, B.D., L.K.K., and M.P.; writing—original draft preparation, V.M. and J.E.M.; writing—review and editing, L.K.K. and S.M.; visualization, B.D. and L.K.K.; supervision, S.M.; project administration, V.M. and L.K.K.; All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: Thanks to Victoria Ackroyd for providing edits to the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Drinkwater, L.E.; Wagoner, P.; Sarrantonio, M. Legume-based cropping systems have reduced carbon and nitrogen losses. *Nature* **1998**, *396*, 262–265. [[CrossRef](#)]
2. Lu, Y.C.; Watkins, B.; Teasdale, J.; Abdul-Baki, A. Cover crops in sustainable food production. *Food Rev. Int.* **2000**, *16*, 121–157. [[CrossRef](#)]
3. Altieri, M.; Nicholls, C. *Biodiversity and Pest Management in Agroecosystems*, 2nd ed.; CRC Press: New York, NY, USA, 2004; ISBN 978-1-56022-923-0.
4. Gardner, J.B.; Drinkwater, L.E. The fate of nitrogen in grain cropping systems: A meta-analysis of 15N field experiments. *Ecol. Appl.* **2009**, *19*, 2167–2184. [[CrossRef](#)] [[PubMed](#)]
5. Gabriel, J.L.; Garrido, A.; Quemada, M. Cover crops effect on farm benefits and nitrate leaching: Linking economic and environmental analysis. *Agric. Syst.* **2013**, *121*, 23–32. [[CrossRef](#)]
6. Schipanski, M.E.; Barbercheck, M.; Douglas, M.R.; Finney, D.M.; Haider, K.; Kaye, J.P.; Kemanian, A.R.; Mortensen, D.A.; Ryan, M.R.; Tooker, J.; et al. A framework for evaluating ecosystem services provided by cover crops in agroecosystems. *Agric. Syst.* **2014**, *125*, 12–22. [[CrossRef](#)]
7. Malone, R.W.; Jaynes, D.B.; Kaspar, T.C.; Thorp, K.R.; Kladvik, E.; Ma, L.; James, D.E.; Singer, J.; Morin, X.K.; Searchinger, T. Cover crops in the upper midwestern United States: Simulated effect on nitrate leaching with artificial drainage. *J. Soil Water Conserv.* **2014**, *69*, 292–305. [[CrossRef](#)]
8. Singer, J.W.; Nusser, S.M.; Alf, C.J. Are cover crops being used in the US corn belt? *J. Soil Water Conserv.* **2007**, *62*, 353–358.
9. Wallander, S. While Crop Rotations Are Common, Cover Crops Remain Rare. In *Amber Waves: The Economics of Food, Farming, Natural Resources, and Rural America*; United States Department of Agriculture (USDA) Economic Research Service (ERS): Washington, DC, USA, 2013.
10. CTIC. *Report of the 2016–17 National Cover Crop Survey*; Joint publication of the Conservation Technology Information Center (CTIC): West Lafayette, IN, USA; The North Central Region Sustainable Agriculture Research and Education (SARE) Program: St. Paul, MN, USA; The American Seed Trade Association: Alexandria, VA, USA, 2017.
11. Wilke, B.J.; Snapp, S.S. Winter cover crops for local ecosystems: Linking plant traits and ecosystem function. *J. Sci. Food Agric.* **2008**, *88*, 551–557. [[CrossRef](#)]
12. Wayman, S.; Kucek, L.K.; Mirsky, S.B.; Ackroyd, V.; Cordeau, S.; Ryan, M.R. Organic and conventional farmers differ in their perspectives on cover crop use and breeding. *Renew. Agric. Food Syst.* **2017**, *32*, 376–385. [[CrossRef](#)]
13. Duvick, D.N. Genetic Contributions to Yield Gains of U.S. Hybrid Maize, 1930 to 1980. In *Genetic Contributions to Yield Gains of Five Major Crop Plants*; John Wiley and Sons, Ltd.: Hoboken, NJ, USA, 1984; pp. 15–47. ISBN 978-0-89118-586-4.

14. Duvick, D.N. Plant breeding: Past achievements and expectations for the future. *Econ. Bot.* **1986**, *40*, 289–297. [[CrossRef](#)]
15. Meredith, W.R.; Bridge, R.R. Genetic Contributions to Yield Changes in Upland Cotton. In *Genetic Contributions to Yield Gains of Five Major Crop Plants*; John Wiley and Sons, Ltd.: Hoboken, NJ, USA, 1984; pp. 75–87. ISBN 978-0-89118-586-4.
16. Miller, F.R.; Kebede, Y. Genetic Contributions to Yield Gains in Sorghum, 1950 to 1980. In *Genetic Contributions to Yield Gains of Five Major Crop Plants*; John Wiley and Sons, Ltd.: Hoboken, NJ, USA, 1984; pp. 1–14. ISBN 978-0-89118-586-4.
17. Schmidt, J.W. Genetic Contributions to Yield Gains in Wheat. In *Genetic Contributions to Yield Gains of Five Major Crop Plants*; John Wiley and Sons, Ltd.: Hoboken, NJ, USA, 1984; pp. 89–101. ISBN 978-0-89118-586-4.
18. Specht, J.E.; Williams, J.H. Contribution of Genetic Technology to Soybean Productivity—Retrospect and Prospect. In *Genetic Contributions to Yield Gains of Five Major Crop Plants*; John Wiley and Sons, Ltd.: Hoboken, NJ, USA, 1984; pp. 49–74. ISBN 978-0-89118-586-4.
19. Russell, G.E. (Ed.) *Progress in Plant Breeding—1*; Butterworth-Heinemann: Oxford, UK, 1985.
20. Brummer, E.C.; Barber, W.T.; Collier, S.M.; Cox, T.S.; Johnson, R.; Murray, S.C.; Olsen, R.T.; Pratt, R.C.; Thro, A.M. Plant breeding for harmony between agriculture and the environment. *Front. Ecol. Environ.* **2011**, *9*, 561–568. [[CrossRef](#)]
21. Yeater, K.M.; Bollero, G.A.; Bullock, D.; Rayburn, A.L.; Rodriguez-Zas, S.L. Assessment of Genetic Variation in Hairy Vetch Using Canonical Discriminant Analysis. *Crop Sci.* **2004**, *44*, 185–189. [[CrossRef](#)]
22. Maul, J.; Mirsky, S.; Emche, S.; Devine, T. Evaluating a Germplasm Collection of the Cover Crop Hairy Vetch for Use in Sustainable Farming Systems. *Crop Sci. Madison* **2011**, *51*, 2615–2625. [[CrossRef](#)]
23. Clark, A. (Ed.) *Managing Cover Crops Profitably*, 3rd ed.; SARE Outreach: College Park, MD, USA, 2007; ISBN 978-1-888626-12-4.
24. Steiner, J.; Piccioni, E.; Falcinelli, M.; Liston, A. Germplasm Diversity among Cultivars and the NPGS Crimson Clover Collection. *Crop Sci.* **1998**, *38*, 263–271. [[CrossRef](#)]
25. Mothapo, N.V.; Grossman, J.M.; Maul, J.E.; Shi, W.; Isleib, T. Genetic diversity of resident soil rhizobia isolated from nodules of distinct hairy vetch (*Vicia villosa* Roth) genotypes. *Appl. Soil Ecol.* **2013**, *64*, 201–213. [[CrossRef](#)]
26. Mothapo, N.V.; Grossman, J.M.; Sooksa-nguan, T.; Maul, J.; Bräuer, S.L.; Shi, W. Cropping history affects nodulation and symbiotic efficiency of distinct hairy vetch (*Vicia villosa* Roth.) genotypes with resident soil rhizobia. *Biol. Fertil. Soils* **2013**, *49*, 871–879. [[CrossRef](#)]
27. Baxter, L.L.; Grey, T.L.; Tucker, J.J.; Hancock, D.W. Optimizing Temperature Requirements for Clover Seed Germination. *Agrosyst. Geosci. Environ.* **2019**, *2*, 180059. [[CrossRef](#)]
28. Shearer, G.; Kohl, D.H. N₂-Fixation in Field Settings: Estimations Based on Natural ¹⁵N Abundance. *Funct. Plant Biol.* **1986**, *13*, 699–756. [[CrossRef](#)]
29. Pinheiro, J.; Bates, D.; DebRoy, S.; Sarkar, D. NLME: Linear and Nonlinear Mixed Effects Models. R Package Version 2019. Available online: https://www.researchgate.net/publication/303803175_Nlme_Linear_and_Nonlinear_Mixed_Effects_Models (accessed on 30 January 2019).
30. Li, D.; Liu, C.-M.; Luo, R.; Sadakane, K.; Lam, T.-W. MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* **2015**, *31*, 1674–1676. [[CrossRef](#)]
31. Aziz, R.K.; Bartels, D.; Best, A.A.; DeJongh, M.; Disz, T.; Edwards, R.A.; Formsma, K.; Gerdes, S.; Glass, E.M.; Kubal, M.; et al. The RAST Server: Rapid Annotations using Subsystems Technology. *BMC Genom.* **2008**, *9*, 75. [[CrossRef](#)]
32. Overbeek, R.; Olson, R.; Pusch, G.; Olsen, G.; Davis, J.; Disz, T.; Edwards, R.; Gerdes, S.; Parrello, B.; Shukla, M.; et al. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Res.* **2013**, *42*, D206–D214. [[CrossRef](#)] [[PubMed](#)]
33. Brettin, T.; Davis, J.J.; Disz, T.; Edwards, R.A.; Gerdes, S.; Olsen, G.J.; Olson, R.; Overbeek, R.; Parrello, B.; Pusch, G.D.; et al. RASTtk: A modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Sci. Rep.* **2015**, *5*, 8365. [[CrossRef](#)] [[PubMed](#)]
34. Knight, W.E. Crimson Clover. In *Clover Science and Technology*; John Wiley and Sons, Ltd.: Hoboken, NJ, USA, 1985; pp. 491–502. ISBN 978-0-89118-218-4.

35. Williams, W.A.; Elliott, J.R. Ecological Significance of Seed Coat Impermeability to Moisture in Crimson, Subterranean and Rose Clovers in a Mediterranean-Type Climate. *Ecology* **1960**, *41*, 733–742. [[CrossRef](#)]
36. Bennett, H.W. The Effectiveness of Selection for the Hard Seeded Character in Crimson Clover1. *Agron. J.* **1959**, *51*, 15–16. [[CrossRef](#)]
37. Fowler, D. Selection for Winterhardiness in Wheat. II. Variation within Field Trials. *Crop Sci.* **1979**, *19*, 773–775. [[CrossRef](#)]
38. Gusta, L.V.; Wisniewski, M. Understanding plant cold hardiness: An opinion. *Physiol. Plant.* **2013**, *147*, 4–14. [[CrossRef](#)]
39. Nair, R.M.; Craig, A.D.; Rowe, T.D.; Biggins, S.R.; Hunt, C.H. Genetic variability and heritability estimates for hardseededness and flowering in balansa clover (*Trifolium michelianum* Savi) populations. *Euphytica* **2004**, *138*, 197–203. [[CrossRef](#)]
40. Salisbury, P.A.; Aitken, Y.; Halloran, G.M. Genetic control of flowering time and its component processes in subterranean clover (*Trifolium subterraneum* L.). *Euphytica* **1987**, *36*, 887–902. [[CrossRef](#)]
41. Abberton, M.T.; Marshall, A.H. Progress in breeding perennial clovers for temperate agriculture. *J. Agric. Sci.* **2005**, *143*, 117–135. [[CrossRef](#)]
42. Gaby, J.C.; Buckley, D.H. A Comprehensive Evaluation of PCR Primers to Amplify the nifH Gene of Nitrogenase. *PLoS ONE* **2012**, *7*, e42149. [[CrossRef](#)]
43. Smith, C.J.; Nedwell, D.B.; Dong, L.F.; Osborn, A.M. Diversity and Abundance of Nitrate Reductase Genes (narG and napA), Nitrite Reductase Genes (nirS and nrfA), and Their Transcripts in Estuarine Sediments. *Appl. Environ. Microbiol.* **2007**, *73*, 3612–3622. [[CrossRef](#)] [[PubMed](#)]
44. Yamazaki, M.; Thorne, L.; Mikolajczak, M.; Armentrout, R.W.; Pollock, T.J. Linkage of genes essential for synthesis of a polysaccharide capsule in *Sphingomonas* strain S88. *J. Bacteriol.* **1996**, *178*, 2676–2687. [[CrossRef](#)] [[PubMed](#)]
45. Bozsoki, Z.; Cheng, J.; Feng, F.; Gysel, K.; Vinther, M.; Andersen, K.R.; Oldroyd, G.; Blaise, M.; Radutoiu, S.; Stougaard, J. Receptor-mediated chitin perception in legume roots is functionally separable from Nod factor perception. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, E8118–E8127. [[CrossRef](#)] [[PubMed](#)]
46. Yan, Y.; Yang, J.; Dou, Y.; Chen, M.; Ping, S.; Peng, J.; Lu, W.; Zhang, W.; Yao, Z.; Li, H.; et al. Nitrogen fixation island and rhizosphere competence traits in the genome of root-associated *Pseudomonas stutzeri* A1501. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 7564–7569. [[CrossRef](#)]
47. Di Cesare, A.; Cabello-Yeves, P.J.; Christmas, N.A.M.; Sánchez-Baracaldo, P.; Salcher, M.M.; Callieri, C. Genome analysis of the freshwater planktonic *Vulcanococcus limneticus* sp. nov. reveals horizontal transfer of nitrogenase operon and alternative pathways of nitrogen utilization. *BMC Genom.* **2018**, *19*, 259. [[CrossRef](#)]
48. Huang, T.-C.; Lin, R.-F.; Chu, M.-K.; Chen, H.-M. Organization and expression of nitrogen-fixation genes in the aerobic nitrogen-fixing unicellular cyanobacterium *Synechococcus* sp. strain RF-1. *Microbiol. Read.* **1999**, *145*, 743–753. [[CrossRef](#)]
49. Liu, D.; Liberton, M.; Yu, J.; Pakrasi, H.B.; Bhattacharyya-Pakrasi, M. Engineering Nitrogen Fixation Activity in an Oxygenic Phototroph. *mBio* **2018**, *9*. [[CrossRef](#)]
50. Riday, H. Correlations between visual biomass scores and forage yield in space planted red clover (*Trifolium pratense* L.) breeding nurseries. *Euphytica Dordr.* **2009**, *170*, 339–345. [[CrossRef](#)]
51. Riday, H.; Brummer, E.C. Forage Yield Heterosis in Alfalfa. *Crop Sci.* **2002**, *42*, 716–723. [[CrossRef](#)]
52. Terpolilli, J.; Rui, T.; Yates, R.; Howieson, J.; Poole, P.; Munk, C.; Tapia, R.; Han, C.; Markowitz, V.; Tatiparthi, R.; et al. Genome sequence of *Rhizobium leguminosarum* bv trifolii strain WSM1689, the microsymbiont of the one flowered clover *Trifolium uniflorum*. *Stand. Genom. Sci.* **2014**, *9*, 527. [[CrossRef](#)]
53. Mcewan, C.E.A.; Gatherer, D.; Mcewan, N.R. Nitrogen-Fixing Aerobic Bacteria have Higher Genomic GC Content than Non-Fixing Species within the Same Genus. *Hereditas* **1998**, *128*, 173–178. [[CrossRef](#)] [[PubMed](#)]
54. Provorov, N.A.; Andronov, E.E.; Onishchuk, O.P. Forms of natural selection controlling the genomic evolution in nodule bacteria. *Russ. J. Genet.* **2017**, *53*, 411–419. [[CrossRef](#)]
55. Muresu, R.; Porceddu, A.; Sulas, L.; Squartini, A. Nodule-associated microbiome diversity in wild populations of *Sulla coronaria* reveals clues on the relative importance of culturable rhizobial symbionts and co-infecting endophytes. *Microbiol. Res.* **2019**, *221*, 10–14. [[CrossRef](#)]
56. Caetano-Anollés, G.; Gresshoff, P.M. Plant genetic control of nodulation. *Annu. Rev. Microbiol.* **1991**, *45*, 345–382. [[CrossRef](#)] [[PubMed](#)]

57. Hirsch, S.; Kim, J.; Muñoz, A.; Heckmann, A.B.; Downie, J.A.; Oldroyd, G.E.D. GRAS Proteins Form a DNA Binding Complex to Induce Gene Expression during Nodulation Signaling in *Medicago truncatula*. *Plant Cell* **2009**, *21*, 545–557. [[CrossRef](#)] [[PubMed](#)]
58. Walker, S.A.; Viprey, V.; Downie, J.A. Dissection of nodulation signaling using pea mutants defective for calcium spiking induced by Nod factors and chitin oligomers. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 13413–13418. [[CrossRef](#)] [[PubMed](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).