|  |  |
| --- | --- |
| **Reporting Period** | Final Report   |
| **Proposal** | [1248 Novel Yield Genes from Cultivated and Wild Japanese Soybean: Gene Cloning and Introgression into Elite Soybean Breeding Lines (Year 1 of 3)](http://moss.unitedsoybean.org/Lists/Proposals/DispForm.aspx?ID=2197&RootFolder=*)   |
| **Committee** | Production   |
| **Target Area** | Supply   |
| **Project Start Date** | 10/1/2010   |
| **Project End Date** | 9/30/2011   |
| **Project Number** | 1248   |
| **Project Status** |  H. Roger Boerma/Zenglu Li, University of Georgia, Athens, GA Thomas E. Carter, USDA-ARS, Raleigh, NCBlair Buckley, Louisiana State University, Bossier City, LA**KPI highlights.**1)    Presence of Yld1 and Yld2 on average (67.6 bu/a compared to 64.2 bu/a) showed over 3 bu/a yield advantage based on 2013 data in a backcrossing population of Boggs x Woodruff. These Yield alleles or genes appear to be the most robust and important found over all USB diversity projects and already have contributed to cultivar releases including N7001, N7002, N8001, and Woodruff. Several new high yielding breeding lines derived from these same cultivars and related materials are now topping the southern USDA southern regional yield trials, and likely carry these yield enhancing alleles from PI 416937.  This is an amazing success story. 2)    SNP marker assays were developed for both Yld1 and Yld2 alleles. 3)    We discovered that approximately 10 genes control plant erectness in wild x domesticated soybean crosses, and that this is the reason that applied breeding with wild soybean has been so difficult in the past. 4)    To overcome the breeding barrier in using the wild soybean for commercial breeding, we developed a novel mega-population approach which produced a large number of adapted breeding lines for commercial breeding from wild soybean hybridizations. 5)    A publication on the Mega-population wild soybean breeding method was generated from this research project. 6)    G07-6012 and G07-6029 derived from N7103 x PI 366122 and having 50% of PI 366122 (*G. soja*) in their pedigrees were approved for release as germplasm by the University of Georgia. G07-6029 is an early maturity group VII line, yielded 88% of its elite parent N7103, while G07-6012 is a late maturity group VII line, yielded 88% of N7103.  7)    A total of 10 wild soybean-derived breeding lines were transferred to four major commercial breeding companies via MTAs in 2012. 8)    A breeding line was developed from the cross of domesticated N7103 soybean x wild soybean PI 366122 that has the yield and maturity of the N7013 parent and derives 20% of its DNA polymorphisms from the wild soybean. This breeding line has been widely distributed to the commercial sector. 9)    Heterosis studies revealed that the wild soybean is a source of yield enhancing alleles for commercial US breeding programs. This is a novel finding which has major implications for commercial breeding and future USB diversity projects.  10) A total of four students were trained through this project (Jake Delheimer and Leah Ruff graduated from NC State Univ. Dave Eickholt a current student at Raleigh; Ben Stewart-Brown, a Ph.D. student at UGA. (Ruff and Eickholt are USB fellows and Ruff is now a Ph.D. student with George Graef in Nebraska). 11) We presented invited seminars on the novel mega-population breeding methods for wild soybean at the University of Missouri, the University of Georgia, and the University of Nebraska.  **Project Status** **1.     Develop recombinant substitution lines that allow cloning of Yld1 and Yld2.** **a)    Set of ‘yield gene NILs’ called ‘X-011828’** Based on the genotypes at Yld1 and Yld2 QTL, we planted 150 out of 280 BC3F4 lines derived from Boggs (3) x Woodruff at 2 locations with 2 replications per location in the summer of 2013 for evaluation of yield and agronomic traits. Notes on the agronomic traits including maturity, plant height, and lodging were taken. 2013 was the first year of yield testing from which we have compiled yield data. Preliminary data from Boggs x Woodruff backcrossing population indicated higher yielding lines associated with the presence of Yld1 and Yld2 on average (67.57 bu/a compared to 64.21 bu/a). Although the difference was not significant, over 3 bu/a yield advantage were observed with both Yld1 and Yld2 QTL. Selected lines based on the yield and marker results will be tested at three locations with 3 reps and 4 row plots in 2014.  There is a distinct possibility that other QTL affecting yield, derived from PI 416937, are present within this Boggs x Woodruff backcross population, so it will be interesting to see if QTL mapping will reveal these QTL. We are preparing DNA from 5 high yielding lines and 5 low yielding lines different allele combinations for genotyping via 50k SNP Infinium chips. When comparing genotyping data between high and low yielding lines at these 50,000 loci, certain haplotypes corresponding to high yield will in theory emerge. If these haplotypes are unique to PI 416937, this will be fascinating to see if other cultivated lines derived from PI 416937 share these same alleles, possibly revealing novel alleles for boosting yield in the relatively narrow gene pool comprising North America’s elite soybean cultivars. The parents, including PI 416937, Woodruff, and Boggs will be sequenced with the help of the Williams 82 reference genome to identify genetic variation in yield QTL within the genome and search for candidate yield genes.   We planted the 280 BC1F3 rows derived from GRAHAM X G03MG-2985 which carried YLD1 and YLD2 alleles as plant rows in the summer of 2013. G03MG-298 was derived from Graham x N96-7031 where N96-7031 is a selection from N7001 × N90-7241 and has 50% of PI416937 in its pedigree. From these 280 plant rows, we have selected and harvested 200 rows. These lines are being genotyped now and selected lines will be yield tested in 2014 summer to validate the yield alleles in different genetic background. Develop SNP markers for selection of two yield genes and confirm the impact of the yield QTL in other genetic backgrounds. Previously, we used SSR markers to identify Yld1 (satt520) and Yld2 (satt333). In this project, based on the Infinium fingerprinting data, we have developed KASP SNP markers that have reliably called alleles identical to SSR marker calls for Yld1 and Yld2 in our Boggs (Yld1/Yld2 absent) x Woodruff (Yld1/Yld2 present) population. For Yld1, the corresponding KASP marker is Gm06\_6952550\_C\_A, located approximately 70 kb away from satt520 in terms of physical distance. For Yld2, the corresponding KASP marker is Gm08\_40734513\_C\_T, located approximately 1.5 Mb from satt333 in terms of physical distance. This is an area of the genome that appears to be relatively lacking in terms of recombination. The next step will be to verify the reliability of these two KASP markers in other populations we have developed which contain Yld1 and Yld2. These new SNP markers, in association with advanced assays, will be used for marker-assisted selection and for tracking yield alleles in our breeding populations. These SNP markers will also be available to breeders from both the public and private sector. Track the marker inheritance across the pedigrees using the Infinium Chip data to develop a novel breeding strategy.  Sometimes alleles show a desirable phenotype in one genetic background or environment but are lost in another. We plan to use RILs derived from G03MG-2985 (Yld1/Yld2 present) by Graham (Yld1/Yld2 absent) to test efficacy of Yld1 and Yld2. Yield testing for this population in this summer. Three breeding populations (G93-2225 x G09PR-54329 (Yld1/Yld2 present); AU02-3104 x G00-3213 (3) RR2Y (Yld1/Yld2 present); NC02-307 x [G00-3213 x RR2Y](Yld1/Yld2 present) having the PI416937 in their pedigrees will be yield-tested in 2014. We plan to use the markers that we developed to track these yield alleles and determine the yield impact of these yield QTL across pedigrees.  We have selected 40-50 high yielding lines derived from PI 416937 or have PI 416937 in the pedigree to be fingerprinted with 50K SNP Infinium Chips. The data will be analyzed at the whole genome level, and compared with soybean ancestral lines and elite North American cultivars to determine which alleles were selected during the breeding process and whether or not any other regions were selected across pedigrees in addition to these two yield QTL. This aggregation of 40-50 lines is currently being compiled primarily by examining pedigrees indicated from Southern States Uniform Soybean Tests in the past 10 years. This is a logical methodology for compiling these 40-50 lines due to the fact that we have trying to uncover yield associated haplotypes and only the highest yielding lines make it to Uniform Soybean Tests. Any yield haplotypes discovered in this analysis will be mapped and verified in future research. Yield associated haplotypes will also be compared to 50k Infinium chip data of North American soybean ancestor lines to see if any yield haplotypes are novel.**b)    Transfer of Asian yield genes to group V and VI maturity.** Thus far, most of our success with the Asian yield genes has centered on maturity group VII cultivars (e.g. Woodruff, N7001, N7002, and N8001).  We attempted to move these genes to elite backgrounds in groups V and VI.  To do this, we crossed N7002 to Clifford.  N7002 traces 25% of its pedigree to Japanese PI 416937.  Clifford is a group V conventional cultivar.  F4-derived lines were yield tested in replicated trials in 2008- 2009 at Kinston NC. The best lines were then tested in 2011-2012 as follows: N07-14221 was tested in USB maturity group V in 2011 and 2012 and in USDA preliminary group V trials in 2012.  Over the 9 environments of the USB trials, N07-14221 matured 3 days later than Clifford and therefore seems to be a very late group V or early group VI breeding line.  In the USB trials where the parent Clifford was also included, this line yielded 7% higher than Clifford (61 vs. 57 bu/a). Combined over all USB and USDA regional trials (23 environments), N07-14221 was 3 days later than the nearest maturing group V check, AG 5606, and yielded 99% of this check. The line was 6 days later than Osage and yielded 96% of this check.  N07-14182, a group VI line, was tested in the USB maturity group 6 trail in 2011 and 2012 over a total of 11 environments, in the USDA preliminary group 6 trials in 2011 (6 environments), and in the USDA Uniform test 6 trials (13 environments) in 2012.  Over the 30 environments, N07-14182 matured on the same day as the high yielding NC-Roy check and yielded 97% of that check. We are screening both lines for the presence of YLD1 and YLD2 marker haplotypes in conjunction with Dr. Li to ensure that these lines carry the QTL.  The increased yield of the group 5 line, N07-14221, over the parental cultivar Clifford is consistent with the hypothesis that yield alleles may be emanating from the Japanese exotic pedigree of N7002 (i.e. PI 416937).  N07-14221 was crossed to Clifford in 2011, progeny advanced *via* SSD, and ~300 F4 plants were harvested in 2013 in NC. These plants were sampled for DNA extraction in the summer of 2013. Working with Drs. Li and Carter, student Dave Eickholt, is preparing to assay these F4 plants for YLD1 and YLD2 alleles using SNP markers and then select plants which have factorial combinations of the YLD1 and YLD 2 alleles. Selected plants will be grown in the field in 2014 as progeny rows to produce the planting seed for replicated yield trials in 2015.**c)    Irradiate Woodruff or its group VII sib and develop a group V mutant as a secondary means of transferring the PI 416937 alleles to group V.** We irradiated approximately 25 lbs of the group VII Woodruff-sib G00-3213 and planted approximately 20 lbs of seed in the field during 2011 at Clayton, NC.  At maturity, these large blocks of irradiated seed were observed for maturity and approximately 220 early maturing plants were identified and harvested individually. In 2012, we planted 250 progeny rows from these single plants. However, none of the progeny rows showed earlier maturity than the original G00-3213. We presume that the early maturity of plants from 2011 resulted from corn root borer damage, which can hasten maturity or outright kill plants when moderate drought occurs during pod filling in sandy fields at Clayton, NC. As a backup plan in 2011 (Plan B) to the single-plant selection described above, we also harvested a pod from all ‘normal-looking’ plants after harvesting early appearing plants. Pods were bulk threshed to obtain approx. 9000 seed. A large block of seed from the ‘plan B’ population was planted at Clayton, NC, in 2012, and observed for maturity. Approximately 50 plants were identified as early maturity and harvested individually. Progeny rows from these 50 plants will be evaluated in 2014 at Clayton, NC.We have prepared seed and will plant in 2014 approximately 20,000 irradiated seed of the group 7 breeding line G00-3213 (a carrier of YLd1 and YLD2) and select early maturing plants for further testing to develop an alternative early maturing source of YLD1 and YLD2. **2.     Verify the proportion of wild soybean (G. soja) in agronomically superior breeding lines from the cross of domesticated x wild soybean, and develop efficient breeding approaches to introgress wild soybean alleles into elite breeding populations.** **New approach- Line Development**Over the past 6 years, two very large replicate F3-populations of G. max cultivar N7103 x soja PI 366122 were subjected to intense selection for ‘max-like appearance’. 200 F-4 derived lines from this selection process were yield tested in 3 to 5 locations from 2008-2010. These breeding lines ranged from 50 to 99% of the yield of appropriate checks. The top 70 lines from this study were retested in 2011 at 3 locations in replicated trials. These same lines were retested at 2 locations in 2012 in replicated trials.  A total of 1600 yield plots were planted in NC related to the max x soja effort in 2012 (not counting regional trials.  A paper on agronomic performance of these lines was accepted by Crop Science in November. This paper is part of Jake Delheimer’s dissertation. Jake presented a poster of this work at the 2011 CSSA meetings in San Antonio and the 2012 Soybean Breeders Workshop in St. Louis in February. SNP Analysis of Breeding LinesDNA for all 200 F4-derived lines for which we have extensive yield data has been assayed for 1536 SNP markers in the USDA Cregan lab using the Illumina bead station protocol. 558 markers were found to be polymorphic.  Preliminary analysis of these genetic marker data, *via* matching coefficients and multidimensional scaling, verifies that each breeding line ranges between 15 and 45% G. soja DNA, averaged over all polymorphic loci.  On average, these lines are about 25% soja DNA, or the equivalent of one backcross to the adapted parent max.  Our positive result from one cycle of breeding with G. soja is clearly unprecedented (and real) and likely to change the way US breeders think about G. soja for applied breeding programs. The SNP marker analysis of the lines is complete. It appears that as many as 10 genomic regions identified via single factor analysis of SNP soja markers may be associated with a positive yield response. The results are consistent with the idea that yield alleles reside in G. soja. We are planning experiments (see below) that will validate or refute the observations.  A paper on DNA analysis of these lines has been drafted by graduate student Jake Delheimer. We have re-analyzed the data using a modified approach to step wise regression, to see if yield enhancing alleles can be identified with greater precision.In 2013, we yield tested over 700 plots to begin the process of verifying the identification of yield enhancing alleles derived from wild soybean.   Number of Genes controlling plant erectness in domesticated x wild crossesWe also have a genetic study in progress to count and potentially identify genes associated with the viny growth habit of wild soybean.  2000 max x soja F2-derived progeny rows (F3 plants in 2010) were grown in replicated trials along with appropriate *G.max* checks and assayed for ability to stand erect at Kinston NC.  Data analysis indicated that 10 to 12 genes govern plant erectness and that erect stand is the product of multiplicative recessive epistasis.  None of the 2000 plant rows were homozygous for erect plants.  No other trait has been documented in soybean as having this level of epistatic interaction. We selected ~100 upright plants from plots in 2010 and conducted a replicated progeny row test for standability in 2011. The replicated study has been evaluated and the most desirable plants harvested.   This study confirms that the frequency of max like F4 plants is very low. None of the F3 plants selected in 2011 produced a progeny row in which all plants appeared max-like. A manuscript describing these results has been drafted and is in internal review. In 2013, we grew 120,000 additional F2 plants from 13 unique max x soja crosses. Observations showed that none of the 120,000 had all the physical attributes of domesticated soybean, and few appeared upright. This observation validates our earlier conclusion regarding the inheritance of plant erectness in domesticated x wild soybean crosses.  These data support the breeding approach of selecting F3–plants from very large populations (outlined by Delheimer) for effective use of wild soybean in applied breeding. **3.     Map and validate yield QTLs from wild soybean.** In earlier work, we identified candidate SNP markers from PI 366122 which appear to be associated with yield.  NILs were developed and Ph.D. candidate Dave Eickholt tested the first sets of these NILS for seed yield in 3 replications at 2 locations in 2013. A total of 90 NILs segregating at 4 putative yield loci were compared along with appropriate checks. Yield data have been summarized, and association with SNP marker alleles is in process. In our most exciting work, we have continued to examine F2 heterosis derived from the mating of breeding lines derived from N7103 x soja PI 366122 back to the adapted parent N7103. We examined F2 heterosis derived from 19 breeding lines at 4 locations in 2013 and found two of the 19 contributed significantly (0.05) to midparent F2 heterosis (10%).  We tested these same 19 populations for heterosis in 2013 and found much greater overall levels of heterosis than in 2012. All but 2 of the 19 populations exhibited a midparent heterosis numerically greater than zero. Average midparent heterosis was 7% and 10 of the 19 populations had 10% or greater midparent heterosis. The increased heterosis may have been due to the very wet year and minimal stress.  The populations exhibiting greatest heterosis for 2013 and 2014 are summarized in the table below.  The results provide strong evidence for the existence of yield enhancing alleles in soja  PI 366122 (Table 1)Table 1. F2 heterosis for Seed Yield in 19 populations over 4 locations in 2012 & 2 locations in 2013.  Populations are derived from lines with 50% soja pedigree crossed with the original max parent.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Female Parent | Male Parent | **2012 MID        Parent     heterosis %**  | **2012****high Parent heterosis %**  | 2013MID Parent heterosis | 2013High parent heterosis | **Ave heterosis over two years** | ***G. soja* SNP Alleles (%)**  |
| **NMS4-44-329** | **N7103** | **+9\*\*** | **+7** | **11** | **7** | **10**  | **37** |
| NMS5-98-3-192 | N7103 | +4 | 0 | 17 | 13 | **10.5**  | 36 |
| NMS5-70-6-129 | N7103 | +6 | -4 | 11 | 5 | **9.5**  | 29 |
| G07-6012 | N7103 | -3 | -9 | 8 | 3 | **2.5** | 21 |

**4.    Develop and release improved germplasm from *G. ma*x x *G. soja* crosses for use in modern breeding.** **Regional Testing**Many ‘max x soja’ lines were entered into the USB Southern Diversity trial and the USDA Southern Regional trials over the past 3 years (all lines tested in such trials at least in 2012). Results from regional southern trials are summarized over locations and years as follows:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Breeding Line | MatGroup | Yield as %N7103  | Yield as % best check | No. of environments | % of 558 SNP markers from soja |
| **NMS4-1-83**  | 7 | 101 | 97 (-2 bu/ac) | 30 | 20 |
| **NMS4-1-77**A  | 7 | 105 | 92 (-4 bu/ac) | 15 | 20 |
| **NMS5-101-2-203**  | 7 | 87 | 79 (-10 bu/ac) | 12 | 36 |
| NMS5-253-1-537 | 7 | 92 | 84 (-8 bu/ac) | 9 | 40 |
| NMS5-112-1-241 | 6 |   | 86 (-8 bu/ac) | 4 | 28 |
| NMS4-52-390 |   |   | 82 (-9 bu/ac) | 20 | 32 |
| NMS5-111-6-240 | 6 |   | 83(-9 bu/ac) | 20 | 32 |
| NMS4-123-651 | 6 |   | 80 (-10 bu/ac) | 7 | 35 |
| NMS4-52-390 | 6 |   | 81 | 24 | 9 |

The following max x soja breeding lines are being tested in the USDA regional trials in 2013: NMS4-19-243 NMS4-6-10 (Maturity Group V), and NMS5-11-1-4 (Maturity Group 7)  The following max x soja breeding lines are being tested in the USB Southern Regional Diversity trials in 2013:  NMS4-6-10 (Maturity Group V), NMS5-112-1-241 (Maturity Group VI), NMS4-1-83, NMS5-11-1-4, NMS5-101-2-203, NMS5-253-1-537 (Maturity Group VII),  Core collection of 15 breeding lines from N7103 x PI366122 that carry all soja derived polymorphisms.  A core collection of 15 breeding lines derived from N7103 x soja PI 366122 were yield tested in two locations in replicated trials in 2013.  All lines yield with 70% of checks and in total carry all ~525 soja SNP polymorphisms identified on the 1536 SNP chip. Our plan is to test these breeding lines in 4 additional environments in 2014 and release them as a set. DNA has been extracted from all lines, plus the parents, and shipped to Perry Cregan’s lab for analysis on the 6K Infinum chip.  Results will be released along with the germplasm and will be used to help clarify the size the genomic regions derived from soja in these lines. These data should facilitate their use in further genetic studies.    Intraline Selection for Yield Enhancement and Simultaneous NIL development to test putative Yield QTL. . Following the pioneering intraline selection approach to yield improvement of Fasoula and Boerma, we were practicing this method in the top 19 yielding F4-derived lines from the *G. max* x *G. soja* yield trials. These lines do not shatter appreciably and almost all carry a phenotypic trait associated with wild soybean or otherwise exhibit the uncommon narrow leaf trait found in the *G. max* parent. Approximately 500 plants of each line were grown in 2011 and 110 individual plants of each line were selected for harvest. Individual plants with obvious visual phenotypic outliers/off-types were culled prior to harvest.  Processing of these single plants has been completed and approximately 1000 plant rows are being grown in 2012 (from a total of 9 F4-derived breeding lines). The remaining plant rows from 10 additional F4-drerived lines will be grown in 2013. SNP marker analysis of F4-derived lines in this project revealed that many were heterogeneous for putative yield genes identified in the Delheimer single factor analysis. Taking advantage of this fortuitous occurrence, these F4-derived lines were prioritized for the intraline selection procedure in 2012.  These lines were successfully harvested in 2012. A new Ph.D. candidate in our lab, Dave Eickholt genotyped approximately 400 sub-lines for contrasting markers.  **5.    Develop NILs for wild/domesticated soybean traits and determine the yield drag associated with each of these domestication traits.** NILs have been developed by a combination of inbreeding and pedigree selection. These Nils were increased in 2009 and yield tested in tests TCLP436 and 437 in 2010 at Clayton NC and in 2011 at Clayton and Kinston NC. In test TCLP437, we evaluated 12 sets of NILS for leaf shape. Preliminary analysis indicates that although leaf shape is multigenic and leaf shape the leaf shape alleles do NOT have appear to have any relation to seed yield.  If this trend holds, then this will be an important finding suggesting that breeders may not need to consider leaf shape in breeding with the wild soybean as it relates to seed yield. In TCLP436, we evaluated 7 sets of NILs for flower color, 3 for seed coat color, 1 pod wall color, and 1 for hilum color.  Preliminary analysis reveals the surprising result that none of these traits appears to be strongly linked with yield.  Tests TCLP436 and 437 and being grown at 2 additional locations each in replicated trials in 2012. We assayed plots from test TCLP 437 at 2 locations for leaf shape in approximately in early September. Data are currently undergoing multi-year analysis. But, the association of morphological traits emanating from wild soybean and seed yield seem minimal indicating that breeders can virtually ignore these traits when selecting progeny derived from the wild soybean.  NOTE: Uploaded by Katie Williams on behalf of Zenglu Li on 3/18/14. |
| **Legacy User ID** |    |
| **Attachments**  |

|  |
| --- |
|  |

 |

|  |
| --- |
| http://moss.unitedsoybean.org/_layouts/images/blank.gif |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|

|  |
| --- |
| Content Type: Project Status Report |

|  |  |  |  |
| --- | --- | --- | --- |
|

|  |
| --- |
| Version: 1.0   |
| Created at 3/18/2014 3:57 PM  by [Katie Williams](http://moss.unitedsoybean.org/_layouts/userdisp.aspx?ID=143)http://moss.unitedsoybean.org/_layouts/images/blank.gifNo presence information   |
| Last modified at 3/18/2014 3:57 PM  by [Katie Williams](http://moss.unitedsoybean.org/_layouts/userdisp.aspx?ID=143)http://moss.unitedsoybean.org/_layouts/images/blank.gif |

 |

 |