STATUS OF THE NATIONAL COTTON GERMPLASM COLLECTION

James Frelichowski
Janna Love
Lori Hinze
Joshua Udall
USDA-ARS
College Station, Texas
Don Jones
Cotton Incorporated
Cary, NC
Jodi Scheffler
USDA-ARS
Stoneville, MS

Abstract

The US National Cotton Germplasm Collection (NCGC) distributes germplasm and associated information of cotton and crop wild relatives to users in the USA and worldwide. It is part of the USDA-ARS National Plant Germplasm System, a collaborative effort to safeguard the genetic resources of agriculture. Annual statistics on key activities such as cotton seed distribution, seed increase, germplasm characterization, acquisition, database development, and research are presented for 2020 and compared with the past 15 years. Legal challenges remain with accessing foreign cotton genetic resources and consequently acquisitions were confined to collection of crop wild relatives in the US and Territories, and with developed cotton germplasm and cultivars whose registration period had expired. Demand for cotton genetic resources remained steady but the rising costs of operations, and static financial resources have reduced to less than a third the capacity to regenerate accessions and generated new backlogs in critical activities. Descriptors and digital images have been gathered on accessions since 2010 to bolster online databases of GRIN-global and CottonGen, and help users refine their seed requests and streamline their research. Characterization and seed increases occurred in several environments, a Cotton Winter Nursery in Liberia, Costa Rica, and summer field plots and greenhouses in College Station, TX. Reductions in seed increases means that demand for cotton seed will outpace the NCGC seed inventories. The trends forecast a widening gap in seed increase and seed distribution and stressed the need for more funding and collaboration to ensure the viability of the NCGC.

Introduction

The curator of the NCGC prepares a report of activities for the full calendar year and sends it to the members of the Cotton Crop Germplasm Committee, National Program Staff members. Beltwide Cotton Conferences are an additional opportunity for these members and key stakeholders in cotton research to review and assess the activities and the status of the NCGC. The report focuses on key activities of seed increase, characterization, distribution, acquisition, and database development. A comparison of annual statistics of seed increase and distribution was performed with the past 15 years. The contents of the NCGC are 9813 accessions of cotton and *Gossypium* seed, housed in humidity controlled, cold storage vaults in College Station, TX and the NCGRP, Fort Collins, CO. It is mostly *G. hirstum* with over 3800 cultivars but also many (~1200) landraces or wild collected cotton. Smaller holdings are *G. barbadense* cultivated and wild collected (~1600), diploid cotton species of *G. herbaceum* and *G. arboreum* (~1800) and the remaining exotic tetraploid and diploid Gossypium species are less than 1000 accessions. Passport data, seed inventories and collection of descriptors and digital images are performed by a curator and technician and posted in online databases of GRIN-Global and Cottongen. The cotton community and public online community requests seed and information through GRIN-Global or direct contact with the NCGC staff. Distributions are made according to inventories and to best match the users' research or educational objectives. Seed increase and characterizations use a tropical Cotton Winter Nursery (CWN) in Costa Rica, Texas A&M field plots, and USDA greenhouses. Cotton Incorporated has a primary manager, V.P. of Cotton Research Don Jones, and a local manager, Alfonso Palafox, for the CWN. Data collection and entry is performed by staff with visits to the nursery or daily observations in local fields and greenhouses. The DBMU, Beltsville, MD and contractor Marty Reisinger assist the staff with GRIN-Global, and CottonGen curator Jing Yu and manager Dorrie Main, assist the NCGC with data and image uploading. Seed acquisitions are from cotton germplasm released to the NCGC from the NCGRP when their protection periods ended and are typically cultivars or germplasm registered with the Journal

of Plant Registration or Plant Variety Protection. Donations are accepted as long as they are free and clear for distributions because the NCGC currently does not enter into Standard Material Transfer Agreements. Conventional cotton germplasm is admitted and distributed. A strip test for common transgenes in cotton was recently applied to incoming cotton germplasm to detect transgenes. Also genetically engineered cotton is currently not incorporated into the NCGC because of insufficient resources to separately curate such germplasm and completely eliminate risks of contamination.

Materials and Methods

Seed Increase and Characterization

The CWN in Costa Rica is normally planted to NCGC accessions that flower only in response to short days and also with conventional seed of Cotton Breeders to advance another generation in winter months. NCGC seed increases and characterization at the CWN (2019-2020 growing season) was 319 accessions (1908 hills, 6 hills each). The breakdown according to species and sub-collection are: 308 *G. hirsutum* TX landraces/racestocks accessions, 8 *G. hirsutum* cotton standards, 3 *G. barbadense* (one, accession GB 713 was direct seeded into carryover nursery). Lori Hinze and Cooperators Benildo Reyes and Kevin Cushman of Texas Tech Univ., sent an additional 233 NCGC accessions to the CWN. The breakdown is: 26 from the *G. barbadense* sub-collection, 96 from the *G. hirsutum* cultivar sub-collection (50 in single hill only), 48 from the *G. hirsutum* landrace/racestock sub-collection, 25 *G. arboreum* and 25 G. herbaceum diploid cotton sub-collections, 6 *G. raimondii* (exotic diploid species) accessions (directly into the carryover nursery). For the NCGC seed increases 9 accessions failed to germinate, 7 only produced one plant, 44 accessions had less than a full or adequate stand. For the collaborator plantings: 90 with full or adequate stands, 27 with less than full stands, 9 with only one plant, and 57 accessions with no plants germinating (mostly with the diploid cotton species). Prior checks of germination maybe necessary to ensure more consistent stands at the CWN.

The CWN is the ideal environment for characterization with morphological descriptors and digital images for the more diverse, wild and photoperiodic cotton accessions. Two trips were made to the CWN by the NCGC curator James Frelichowski and NCGC technician Janna Love to bracket more of the growth stages and to split the workload because diverse germplasms range in their maturation rates. The first trip was from February 2 to 8. Janna also trained one of the CWN workers, Mario, to collect most of the leaves, flowers, bolls, and stems for digital images. Also attending was Jodi Scheffler who assisted Janna after completing her research. The second trip was from March 1 to 7 and additional attendees were Benildo Reyes and Kevin Cushman of Texas Tech University, Lori Hinze, and Josh Udall. After they completed their research, they assisted Janna. A transplant nursery was established in the back of the CWN for late maturity accessions from previous seasons and was also observed for descriptors and images taken. The descriptor scoring and image collection filled in gaps for critical accessions and at least 143 accessions had new descriptors rated. The list of descriptors is online at CottonGen https://www.cottongen.org/data/trait/NCGC_rating_scale and in GRIN-Global via descriptor tab, crop COTTON https://npgsweb.ars-grin.gov/gringlobal/descriptors. Representative leaf, flower, stem, and bolls were tagged from each accession and taken to covered garage at the CWN for photography. Each plant part was positioned on white backed omnigrid underneath high-resolution Canon digital cameras affixed to camera stands. Alfonso helped acquired and setup supplemental lighting to remove shadows and speed up focusing. A mobile demand field tablet with shoulder straps was used for field scoring of descriptors. It was preloaded in Microsoft Excel software with the accessions and plot numbers for rows and columns for each character. Each cell had a specific drop-down menu (enabled by data validation function) of ratings for a trait.

A nearby field from Texas A&M is used each summer to plant accessions of the NCGC for observation and/or seed increase. Typically accessions in need of characterization are first observed and photographed in College Station as a germination check and if photoperiodic (failed to flower in long summer days) scheduled for replanting and seed increase at the CWN. The 2020 summer planting in College Station focused on descriptors, images and critical seed increases of select NCGC accessions that are day neutral flowering. The total planted was 134 accessions and it breaks down to: 117 *G. hirsutum* (8 were standards), 2 *G. barbadense* (standards 3-79 and Pima S-6), 2 *G. herbaceum*, and 4 *G. arboreum*, 3 *G. sturtianum*, 1 *G. nandewarense*, 1 *G. thurberi*, 1 *G. stocksii*, 1 *G. longicalyx*, 1 *G. australe*, and 1 *G. nelsonii*. Self-pollinations were initiated on all but 36 accessions (for images and descriptors only). All accessions flowered except for *G. longicalyx*. Self-pollination involved covering ‘candle stage’ buds with organza mesh bags. The diploid species accessions of *G. sturtianum* and *G. nandewarense* had bags momentarily removed for sib pollination with a small brush. *G. australe* and *G. nelsonii* typically had permanently

closed or cleistogamous flowers. The other exotic diploids are not cross compatible so open pollinated seed was harvested from them. Descriptor ratings were scored with same tablet and another excel sheet of accessions as rows and each trait as columns, and cells with drop down menu to select a rating for that trait. Representative leaf, flower, stem, and bolls were tagged and brought indoors to USDA workspace with camera stands and controlled lighting.

The 2020-2021 season at the CWN began with the USDA in College Station collecting cooperators seed, arranging for the inspection to obtain a phytosanitary certificate, final packaging with lists and official documents, and FedEx shipping to Costa Rica. The procedures perfected by Don Jones and Alfonso Palafox in the previous season, were replicated with APHIS (PCIT online portal), the Texas Dept. of Ag., and the same Costa Rican contacts identified by Alfonso. Seed was all acid delinted, dusted with Captan, and in individual 2.5" x 4.5" envelopes according to each hill of planting at the CWN. Each user packaged their seed into is 18" x 18" x 10" boxes and contents labeled on outside and on an excel sheet. Copies of seed inventories, cover letter from the Curator, Alfonso's import permit, APHIS phytosanitary certificates were placed inside and outside of the boxes for final shipment. The same paperwork and original phyto were FEDex to AG AD GLOBAL SERVICIOS and a scanned pdf of everything was also emailed to them because of strict Costa Rica importing procedures. The first shipment was sent in mid-October with seed of 3 cooperators. The NCGC portion involved 307 accessions. (1842 hills, 6 hills each). Breakdown is: 2 *G. barbadense* (2 standards), 107 existing *G. hirsutum* landrace sub-collection accessions and 63 new accessions from the 2013 collection of wild cotton in Puerto Rico, 127 *G. hirsutum* SA or cultivar sub-collection, 8 *G. hirsutum* standards. A second shipment of 4 cooperators passed inspection and was shipped to the CWN in early November.

In the Materials and Methods section, give enough detail to allow a competent scientist to repeat the experiments. This section may be arranged in any logical manner and may include tables and figures. All text in the main body of the paper should be 10 pt Times New Roman font, black font color, single-spaced, and with full justification. Double-space below the title 'Materials and Methods' and the first paragraph of the section. Double-space between sections and double-space between paragraphs within sections. To enter your text, simply highlight the paragraph here and replace with your material.

Distribution and Acquisition

The NCGC distributed seed and information to users via GRIN-Global shopping carts, or after contacting James or Janna. Distributions in 2020 were 2098 accessions in 119 orders. Most orders originated in the USA (116 orders, 2090 items), with only three (9 items) from abroad. The breakdown of items is: *G. hirsutum* cultivars (827), *G. hirsutum* landrace/racestocks (167), *G. barbadense* (118), diploid cotton (924), exotic tetraploid (13), and diploid species (49). The orders were also broadly characterized by type of users and broken down as: ARS 7 orders, 62 items, University 31 orders, net 626 items (640 req, 1 sub, 15 n/a), Seed company/Seed bank 11 orders, net 1123 items (1135 req, 12 n/a), Farms/related businesses 13 orders, net 90 items (87 req, 9 sub, 6 n/a), home growers 37 orders, net 131 items (137 req, 2 sub, 8 n/a), educators 20 orders, 57 items. The last two represent a growing base of users fueled perhaps by online access through GRIN-Global and staying at home on account of Covid19 and planting crop seeds as a hobby or for educational purposes). Acquisitions were conservative to prevent duplication within the NCGC and because foreign germplasm often comes with legal compliance with international treaties on intellectual property rights and benefit sharing. Therefore, collection trips by the NCGC are currently confined to the U.S. and its territories. The list of recent acquisitions from collection trips or donations are: 7 wild cotton accessions from private land owners, conservations, or townships in South Florida, 1 *G. thurberi* from Perin McNelis at Borderlands Restoration Network (assisted by Karen Williams, PEO, Beltsville, MD), 4 more *G. thurberi* from Mike Cashman at W6 Western Regional Plant Introduction Station, Pullman, WA, (assisted by Brian Irish, PGITRU, Prosser, WA), 8 donated from Warren Conaty of CSIRO, Narrabri, Australia (assisted by Brett Rose, Cotton Seed Distributors Ltd. Wee Waa, NSW Australia), 4 high oleic lines developed by Michael Dowd and collaborators of SSRC, New Orleans, LA, CSRL, Mississippi State, MS. Availability of accessions released to the NCGC from PVP or JPR protection periods was delayed because of recent NCGC policy of initial testing for adventitious presence of transgenes and ensuring increases and distribution of pure self-pollinated seed. Known transgenic accessions were withheld from the NCGC because of resource limitations and inability to handle completely separate fashion to avoid cross contamination.

Database

Cotton morphological descriptors and digital images have been collected at the CWN and local summer field plantings for over a decade and most of the non-critical accessions have been characterized. This information was

formatted and regularly given to Jing Yu, curator of Cottongen, and Dorrie Main, of Washington State University. Jing diligently compiled accessions from the NCGC, and linked them to passport information, descriptor ratings, and digital images. In some cases, accessions were cited to genetic research and even compared to other cotton collections. However, there is a backlog of uploading this information to the USDA site, GRIN-Global. The past year James and Janna were trained to use GRIN-Global by DBMU contractor Marty Reisinger, with emphasis on formatting data from excel for easy uploading online. Janna updated existing NCGC seed inventory and completed this year's orders in GRIN-Global, while James formatted the new standardized set of descriptors and began uploading them. Jing accepts all observations and images from the NCGC and presents them in CottonGen, but James and Janna had to follow a sequence of creating an inventory and then linking observations in GRIN-Global. Public access to GRIN-Global is shown in Figures 5 and 6. Public users should first register, log in, then go to 'Descriptor' tab, then scroll down for COTTON ('Pre2006 Cotton' is the descriptor data uploaded prior to 2006 with old formats and often specific to sub-collection). Figure 7 shows the COTTON descriptors currently uploaded and further subclassified by types. Currently observations matching actual seed inventory are uploaded but Janna will likely finish inventory creation but with a different status (pending, not received, observation only, etc.) so James can load more observations. The standard descriptor panel of 36 characters along with their rating system is shown in Figure 8 and posted online at http://www.cottongen.org/data/trait/NCGC_rating_scale and many have picture standards linked to each character. Examples of such pictures are in Figure 9. To date a set of *G. hirsutum* 'SA' was uploaded in GRIN-Global and can be viewed at https://npgsweb.ars-grin.gov/gringlobal/search. Try a search for PI 540885 and click on 'observations' to see those descriptors for accession Acala Maxxa (SA 2288). To get the most information on the NCGC one should utilize both sites.

Significant progress in characterization was made (80% of the NCGC) by first growing germplasm that is not in critical status and has high germination rates. Even with photoperiodic germplasm only sent to the CWN plants were first grown in Texas to ascertain their germination rates and to collect a set of descriptors and images so only gaps remained to be filled with visits to the CWN. The remainder (20%) has low germination rates and critical seed inventory. Even when seed has been successfully germinated additional challenges are photoperiodic flowering, low yields, special care of odd growth habits as perennial trees or vines that resprout from basal roots, and exerted stigmas that require special care to prevent outcrossing and ensure pure self-pollination.

Results and Discussion

This report is standard procedure for the NCGC curator to submit to the CCGC just prior to the annual Cotton Beltwide Meetings. A recap of 15 years of activity was included in this report to show the status of activities and forecast trends. What follows are a series of figures using data from each of the 15 years of activity in reference to a 15-year average. A review of plantings at the CWN since 2006 was performed to show the importance to the NCGC and changes over time. Yearly and 15-year averages on total accessions planted and portions from each sub-collection were used. Total planting numbers at the CWN since 2006 are shown in Figure 1. NCGC plantings over this period averaged 816 per year and the yearly amount has declined since relocation from Tecoman, Mexico to Liberia, Costa Rica (2015 was the transition year without a crop). The operating expenses and per hill costs have risen going from as little as 9 USD$ to currently about 27 per hill effectively reducing the amount that the NCGC spends annually to the CWN to about one third. Only photoperiodic accessions are being sent to the CWN while day neutral accessions are increased locally in College Station. The NCGC once had the resources to target most accessions on a 10-year seed increase cycle, but now each accession will likely average one seed increases every 20 – 30 years. The proportion of the NCGC according to sub-collection planted at the CWN has varied over this time period. Before 2015 there were more diploid cotton and *G. hirsutum* cultivars planted at the CWN, but since the move to Costa Rica it is primarily *G. hirsutum* landraces 'TX' sub-collection.

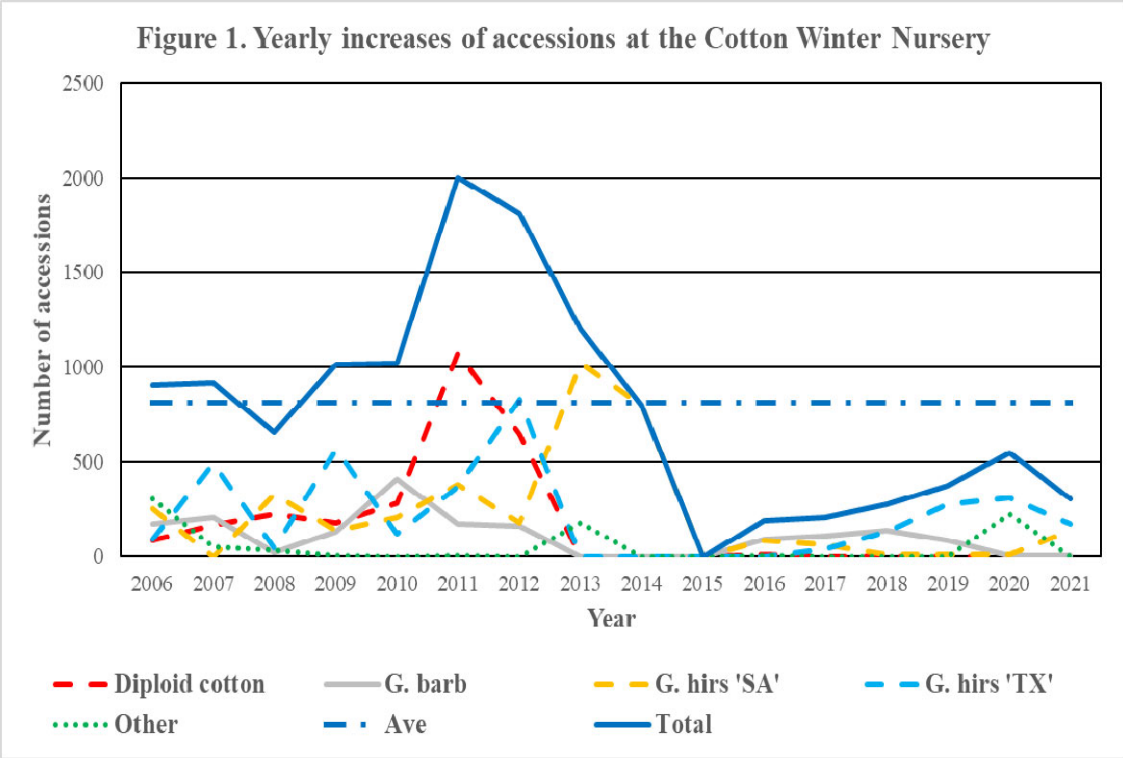

Figure 1. Cotton Winter Nursery plantings of accessions of the NCGC from years of 2006 to 2021.

It is difficult to identify trends in orders for the past 15 years, but it appears to have two sources: regular cotton breeding or research programs or new visitors to GRIN-Global wishing to learn more about cotton and *Gossypium*. Spikes often result from individual large orders following priority research projects (e.g. screening for FOV or virus resistances). Figure 2 shows the total yearly orders over this span and the numbers from domestic and foreign users. Domestic use spiked in 2014 and 2015, and at the same time of relocating the CWN. This put a strain on producing critical seed to match demand.

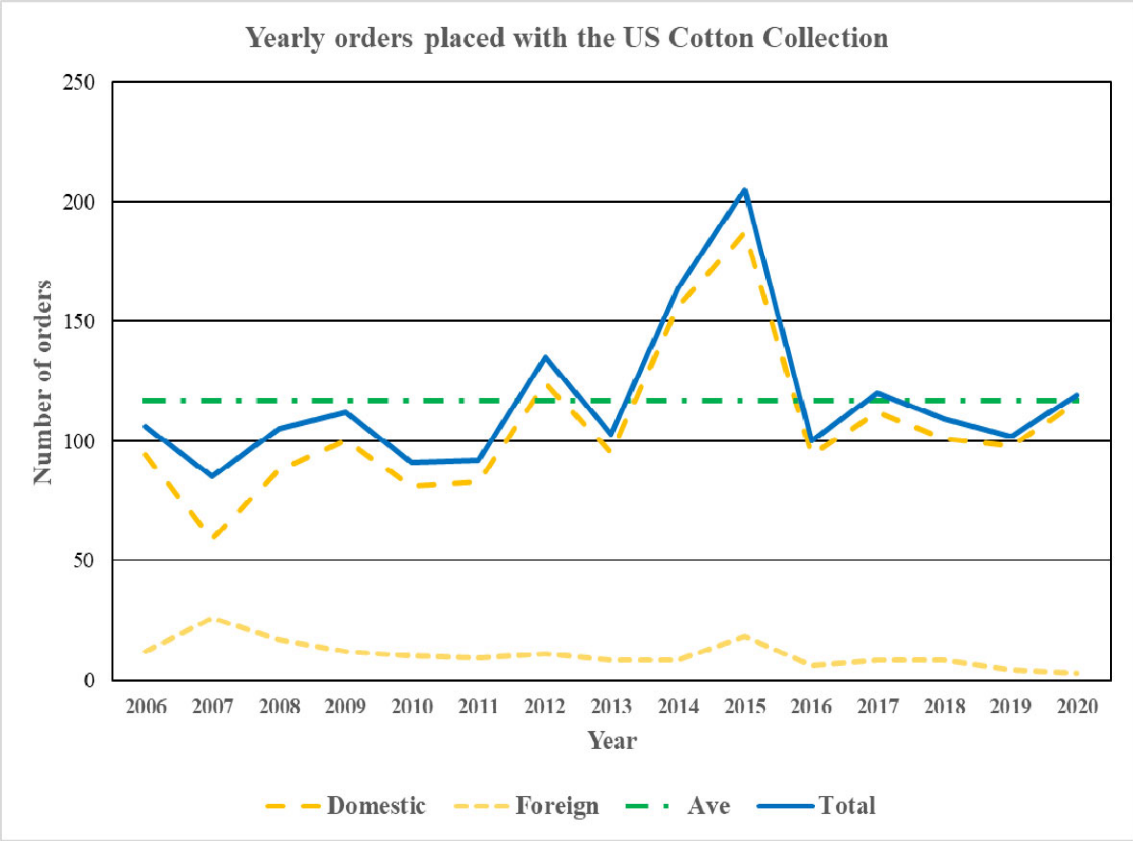

Figure 2. Yearly orders placed with the NCGC from the years 2006 to 2020.

It was difficult to see any meaningful pattern on distributions of items over this time period in Figure 3 other than domestic users demand far more than foreign users, who also have importing hurdles to clear in their countries before we even consider sending them seed.

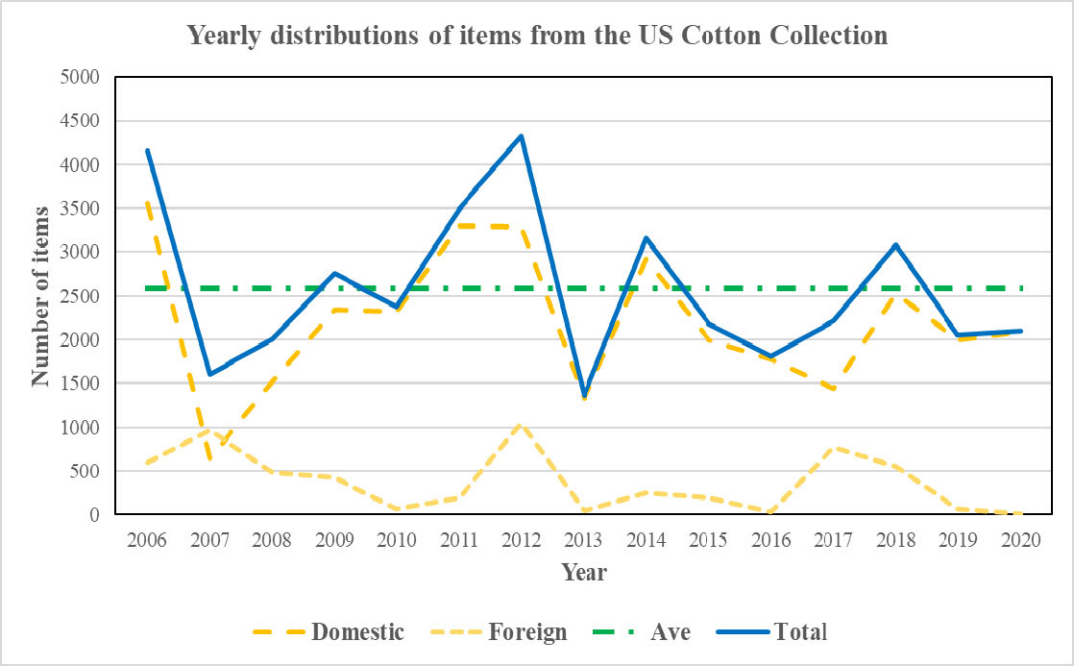

Figure 3. Yearly distributions of items from the NCGC from the years of 2006 to 2020.

Another analysis of the orders is according to sub-collection. Break down of the NCGC by sub-collection is relevant because different species have vastly different characteristics agronomically and particularly with the resultant fiber. The *G. hirsutum* cultivar sub-collection typically includes an inventory prefix of 'SA' because it originated with obsolete cultivar accessions maintained in Stoneville, MS in the Delta Cotton Production Region. *G. hirsutum* landrace/racestock sub-collection has inventory prefix of 'TEX' (or TX) because USDA curators in College Station, TX accumulated wild collected and diverse *G. hirsutum* and needed to distinguish it from cultivars because of different flowering responses and other un-adapted growth characters. *G. barbadense* sub-collection originated with USDA breeders in Sacaton and Maricopa, AZ because of the development of Pima cottons for that region. But this species has different fiber properties, longer growing season and genetic challenges in hybridization with *G. hirsutum*. Asiatic and African diploid cotton species are kept in another sub-collection because while they have fiber and are easily grown along cotton cultivars, the fiber quality is vastly inferior to commercial cotton.

The items sent per year were broken down by sub-collection in Figure 4 and they show a general preference for *G. hirsutum* cultivars, with less for *G. hirsutum* landraces and *G. barbadense*. Recent decline for *G. hirstum* landraces could be combination of high priority breeding with just adapted cotton germplasm and inability to distribute seed of the most critical (lowest inventory) accessions. In some years the demand for diploid cotton was second highest (2014 and 2020) and it was generally from a single research project. Demand for *G. barbadense* from 2015 to 2017 peaked, perhaps sparked by interest in improving disease and nematode resistance or fiber quality. Demand for remaining species remained the lowest and is the most difficult group to grow economically or even as a horticultural plant. Keeping up with demand of the most critical and exotic accessions requires intensive maintenance in USDA greenhouses. That is why attempts are made to grow some outdoors in summer fields of College Station for day neutral accessions or in the carryover nursery at the CWN for photoperiodic accessions.

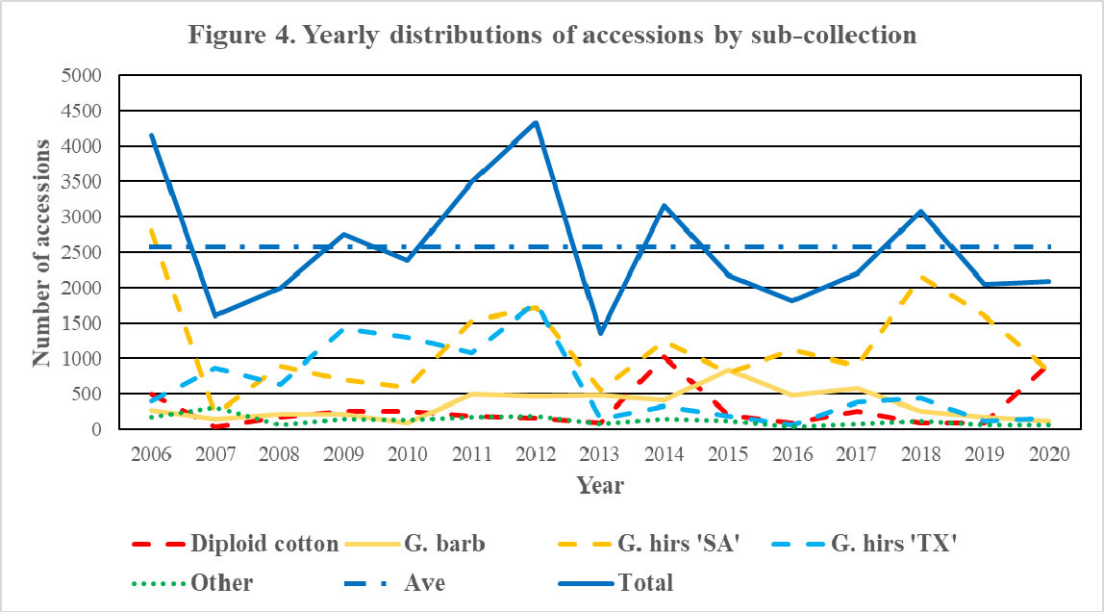

Figure 4. Yearly distributions of items from the NCGC according to sub-collection from 2006 to 2020.

Table 1 summarizes the averages per year in the past 15 years of distributions (and their sub-collections), plantings at the CWN (and sub-collections), and acquisitions of new germplasm. The yearly average for acquisitions is 23 and mostly composed of expired varieties that expand the 'SA' subcollections. Limited additions to the 'TX' and other species were restricted to collection of wild cotton in Puerto Rico (~70 accessions in 2013), South Florida (~5 accessions in 2020) and *G. thurberi* in Arizona (~30 accessions from 2015 and 2017).

Table 1. 15-year averages for accessions of the NCGC distributed, sent for seed increase to the CWN, or new acquisitions. Additional breakdown according to items requested domestically or from foreign user, items in a sub-collection or other germplasm (breeding lines) is listed.

Category	Sub-category	yearly average of accessions
All orders	Total	**2579**
	Domestic	2200
	Foreign	379
	Diploid cotton	292
	Gossypium barbadense	352
	Gossypium hirsutum 'SA'	1180
	Gossypium hirsutum 'TX'	624
	Other species	131
CWN increase	Total	**815**
	Diploid cotton	178
	Gossypium barbadense	113
	Gossypium hirsutum 'SA'	241
	Gossypium hirsutum 'TX'	229
	Other germplasm (lines, stocks)	54
Acquisitions		**23**

Because every order and their requested items were not adjusted for multiple requests by the same user or same accession across different orders, Table 2 was created to summarize number of requests/accessions.

Table 2. Number of accessions according to Frequency of requests.

Times requested	Accessions
once	1289
2 – 10 times	5666
11 – 20 times	581
21 – 30 times	56
31 – 40 times	15
41 – 50 times	7
51 – 60 times	3
60+	1
Total	7618

Table 2 shows 1289 accessions only requested once in the last 15 years. At the other extreme is one that had over 60 requests (NCGC ID is GB 713). The total is 7618 accessions which means at least a quarter of the NCGC has not been distributed. Table 3 lists the top 36 accessions with the most requests in the past 15 years. GB 713 tops the list and is likely the result of positive press as a source of resistance to nematodes and the enigmatic name of 'Inca Cotton'. But the seed inventory is critical, and most requests were sent a substitute (e.g. Pima S-6), which is more suitable for novice home growers and educators. Certain traits such as red leaves/stems, okra leaf, green lint, and an interest in high quality fiber in *G. barbadense* appear to be in high demand. Several species are highly requested perhaps for vital traits such as disease or nematode resistance, confirmed at least in *G. longicalyx*. *G. raimondii* gets many requests possibly because of scientific press as the progenitor of the D subgenome of tetraploid cotton and recently sequenced.

Table 3. Most requested accessions of the NCGC in the past 15 years

Inv Prefix	Inv Num	Taxon	PI	Name	Requests
GB	713	*G. barbadense*	608139	INCA COTTON	63
SA	1213	*G. hirsutum*	529278	COKER 312	60
GB	1030	*G. barbadense*	608346	PIMA S-6	55

GB	1023	*G. barbadense*	560140	PIMA S-7	54
SA	97	*G. hirsutum*	528504	SUPER OKRA	48
SA	2269	*G. hirsutum*	607172	TM 1	47
SA	24	*G. hirsutum*	528441	RED OKRA GREEN CLEAN	46
GB	1381	*G. barbadense*	630149	PIMA OKRA LEAF	43
SA	3	*G. hirsutum*	528420	ARKANSAS GREEN LINT	42
SA	1184	*G. hirsutum*	529249	COKER 310	42
SA	1512	*G. hirsutum*	529566	DELTAPINE 50	41
SA	1475	*G. hirsutum*	529529	DELTAPINE 90	40
SA	1611	*G. hirsutum*	606809	DES 119	39
SA	240	*G. hirsutum*	528608	ACALA RED OKRA LEAF	38
SA	686	*G. hirsutum*	528891	DWARF I OKRA CLEAN	38
D_5	2	*G. raimondii*	530899		38
SA	25	*G. hirsutum*	528442	INTENSE RED GREEN LINT	36
SA	2288	*G. hirsutum*	540885	ACALA MAXXA	36
F_1	3	*G. longicalyx*	530988		36
F_1	1	*G. longicalyx*	530986		35
D_5	31	*G. raimondii*	530928		34
SA	2580	*G. hirsutum*	612326	ACALA 1517-99	34
D_5	1	*G. raimondii*	530898		34
SA	2291	*G. hirsutum*	607174	DELTAPINE 61	32
SA	1668	*G. hirsutum*	606814	PAYMASTER HS 26	31
SA	1723	*G. hirsutum*	547084	LA-887	31
SA	2903	*G. hirsutum*	612959	GA 161	31
GB	679	*G. barbadense*	608115	BLEAK HALL SEA ISLAND	30
SA	9	*G. hirsutum*	528426	INTENSE RED OKRA CLEAN	30
SA	739	*G. hirsutum*	528930	AK-DJURA RED-OKRA-NANKEEN	30
D_1	1	*G. thurberi*	530765		29
SA	398	*G. hirsutum*	528720	OKRA ROUND	29
SA	1186	*G. hirsutum*	529251	DELTAPINE 16	29
SA	1198	*G. hirsutum*	529263	STONEVILLE 7A OKRA	29
SA	1598	*G. hirsutum*	606807	RRB2-10 (SUB OKRA)	29
SA	1595	*G. hirsutum*	606805	PD 1	28

Two databases were used to upload NCGC data. Seed inventories and descriptors were uploaded to GRIN-Global and descriptors and full-sized digital images were sent to Jing Yu for CottonGen. With training from Marty Reisinger, contractor with DBMU in Beltsville, MD, significant progress was made with updating GRIN-Global inventories and descriptors. Janna updated the entire seed inventory of the NCGC, and James formatted and began uploading descriptors online. Because James and Janna have cycled through portions of the collection, using different environments, it was decided to pair each observation to each planting, thus creating a new inventory each time in GRIN-Global. To view these descriptors in GRIN-Global visit their site, click on Descriptor tab (Figure 5), and select COTTON (Figure 6) in crops. Cotton pre2006 is archived data and often has ratings specific to each sub-collection, but COTTON has descriptors with ratings applied to all the sub-collections and *Gossypium* species. This later approach should make it easier to compare diversity among all accessions.

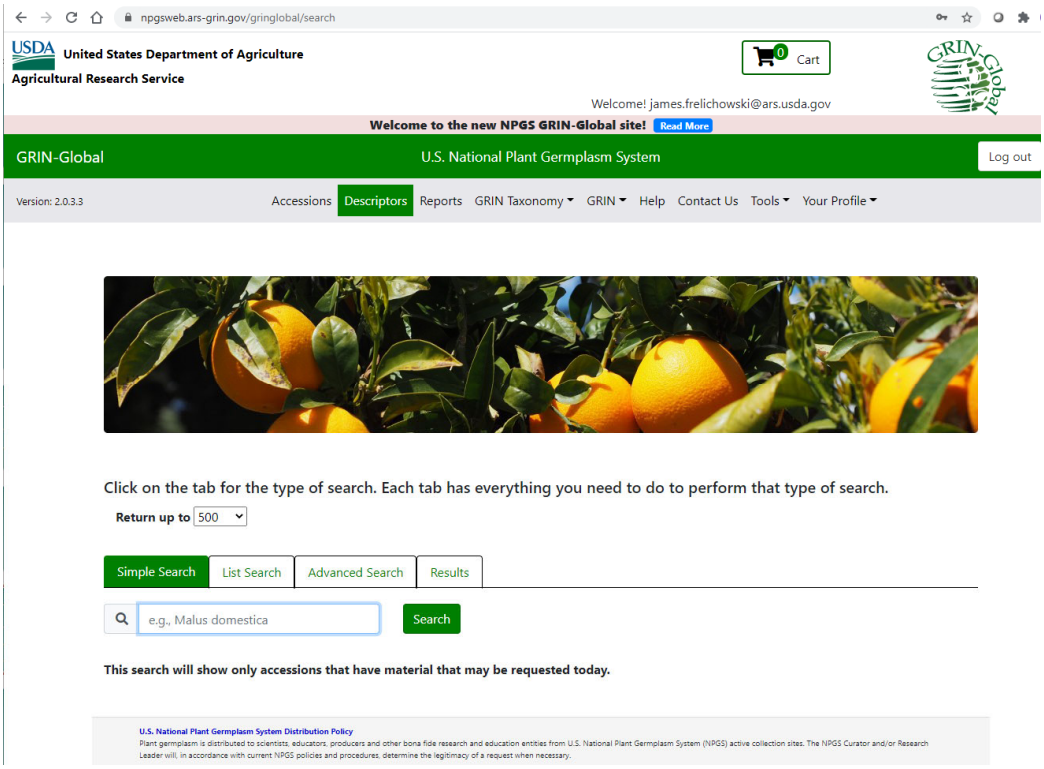

Figure 5. GRIN-Global webpage for public users, emphasizing the 'descriptors' tab.

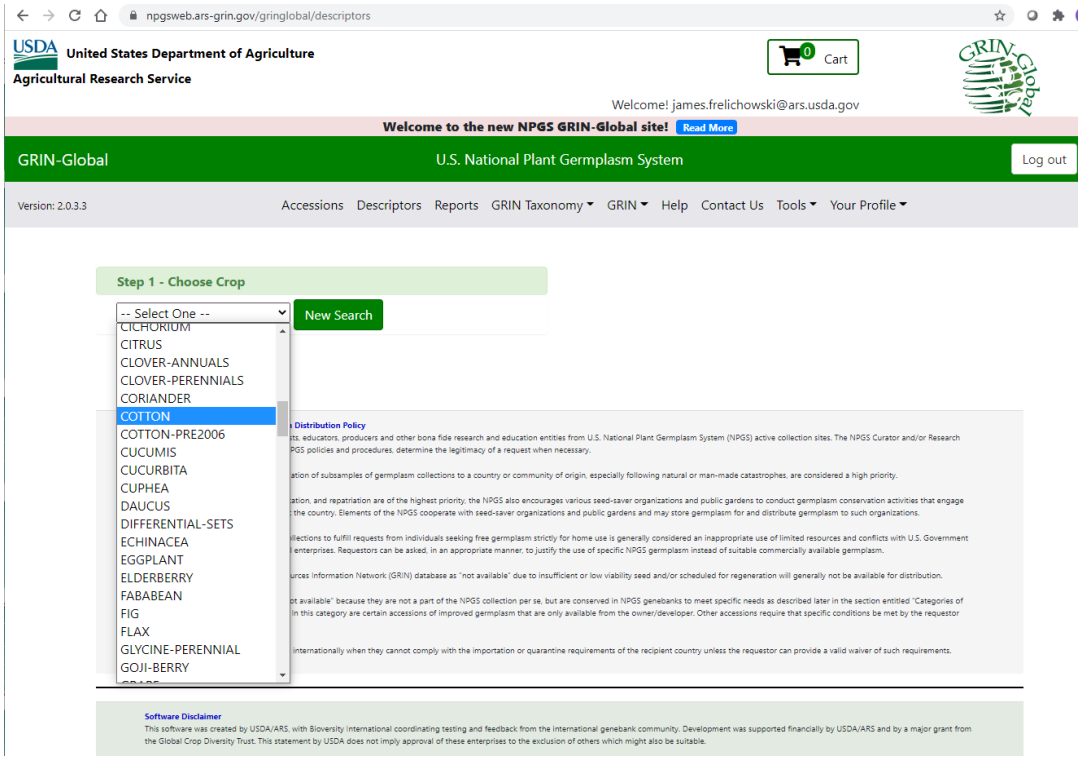

Figure 6. GRIN-Global webpage for users, in descriptors page, scrolling down to select 'COTTON' where the new, standardized morphological descriptors of the NCGC are uploaded

The progress to date on descriptors is with a set of *G. hirsutum* cultivars had their inventory and descriptors updated in GRIN and data is accessible by viewing descriptors and clicking on each trait (Figure 7) or typing in a completed accession such as PI 540885, better known as 'Acala Maxxa', and is given the NCGC inventory ID of SA 2288. The drawback of the current approach in GRIN-Global is if we wait until seed is harvested then it will delay the addition of data to GRIN-Global because entries follow the creation of new seed inventory.

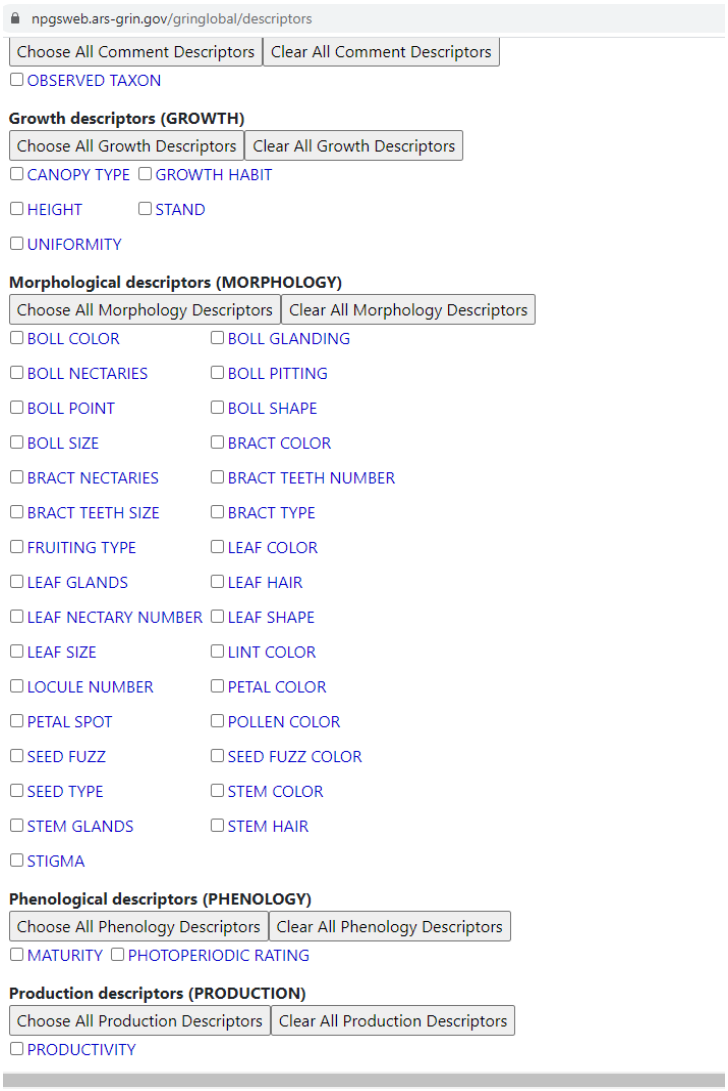

Figure 7. GRIN-Global descriptors currently in the COTTON collection database. Further selection of each trait will show summaries in the database, along with individual accessions and their rating.

Our new approach will likely involve creating inventory during planting but assigning a status of 'pending' or 'observation only' so that descriptor uploads can occur more readily, and the status changed when seed is finally counted and stored. It may create multiple descriptor ratings for some accessions, but it more realistically portrays the phenotypic diversity. A set of 10 standards have been planted each time that may illustrate the variation particular to each environment or year. Fortunately, Jing Yu and Dorrie Main already designed a cotton specific website CottonGen so that genetic, genomic, and germplasm data could be posted for the US and worldwide cotton community. Their expertise allowed for rapid posting online of descriptors and images while the NGCG works to format them for GRIN-Global. They already listed the whole set of descriptors and ratings online (Figure 8) but in

GRIN-Global only the descriptors and ratings that were uploaded are visible, not the full reference set. Online links between the digital images and descriptors were created in CottonGen (Figure 9). It is recommended to utilize both sites for learning more about the NCGC.

cottongen.org/data/trait/NCGC_rating_scale

Descriptor	Description Code or Rating Scale									
	0	1	2	3	4	5	6	7	8	9
Boll Color*		green	dark green	light green	red	sun red				seg/off type
Boll Glanding	glandless	medium	light	heavy						seg/off type
Boll Nectaries	absent	present	reduced	inactive						seg/off type
Boll Pitting		smooth	light pitted	pitted	very pitted					seg/off type
Boll Point		Moderately pointed	pointed	blunt						seg/off type
Boll Shape		Oval	round	cone	cone-oval					seg/off type
Boll Size		medium	large	small	extra small					seg/off type
Bract Color		green	red	sun red						seg/off type
Bract Teeth Number		medium	many	few						seg/off type
Bract Teeth Size	none	medium	small	large						seg/off type
Bract Type		normal	frego	flared	recurved					seg/off type
Fruiting Type		normal	cluster							seg/off type
Growth Habit		normal	spreading	prostrate	pyramid	stovepipe				seg/off type
Leaf Canopy		typical	open	dense	compact					seg/off type
Leaf Color		green	red	dark red						seg/off type
Leaf Glands	glandless	medium	light	heavy						seg/off type
Leaf Hairs		none	few	moderate	hairy	very hairy	pilose			seg/off type
Leaf Nectaries	absent	present one of following: one, main vein(1-1), two(1-2), three(1-3), four(1-4), five(1-5)					reduced (2)			seg/off type
Leaf Shape		normal	okra	sub okra	super okra	laciniate	ovate	cordate		seg/off type
Leaf Size		medium	small	large	extra small					seg/off type
Lint Color		white	cream	brown	green	tan	rust	off white		seg/off type
Locule #		one	two	three	four	five	> five			seg/off type
Maturity		not flowering	flowering	< 1/2 open	> 1/2 open	complete				seg/off type
Petal Color		cream	yellow	light yellow	red	white	light blue	golden	pink	seg/off type
Petal Spot	none	light	medium	heavy						seg/off type
Photoperiodic Rating	no squares	squares	flowers	bolls	open bolls					seg/off type
Plant height (m)	actual measurement									
Pollen Color		yellow	cream	--	dark yellow	orange				seg/off type
Productivity	none	< 12/plant	13-24/plant	25-36/plant	37-48/plant	>48/plant				seg/off type
Seed Fuzz	none	medium	high	sparse	tufted					seg/off type
Seed Fuzz Color	lintless	white	cream	brown	green	tan	rust	off white		seg/off type
Seed Type		free	semi kidney	kidney						seg/off type
Stem Color		green	sun red	red						seg/off type
Stem Glands	glandless	medium	light	heavy						seg/off type
Stem Hair		none	few	moderate	hairy	very hairy	pilose			seg/off type
Stigma		normal	protruding	extreme protruding	short buried					seg/off type

Figure 8. CottonGen database of descriptors of the NCGC. The full range of ratings and shorthand description of each rating for each trait is listed. Descriptors in green have images for the ratings.

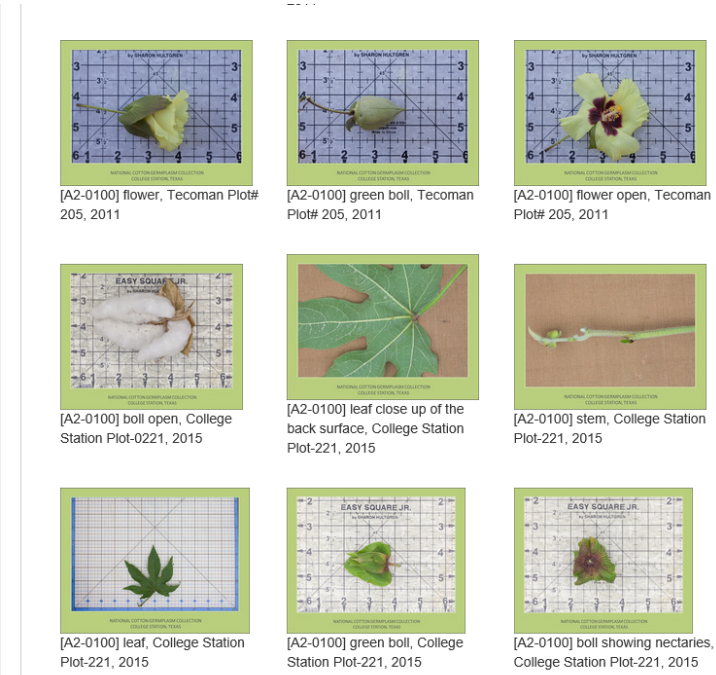

Figure 9. CottonGen images for accession PI 529728 (NCGC inventory ID as A2 100) of the NCGC. Leaf front, back, flower, boll, boll base, opened boll and stem tip are the standard images collected on accessions grown at the CWN or College Station fields.

Summary

The CWN was highly successful because of timely planting and care by the CWN crew. Characterization was successful because of two trips and assistance from other users of the CWN. This 2020-2021 season started on time because of shared experiences on how to prepare and send seed to clear Costa Rica Customs. Success at the CWN is essential to restoring regular seed increases of photoperiodic accessions and to replenish the most critical accessions that still constitute nearly a quarter of the NCGC.

Distribution of germplasm remains high and places a strain on limited seed stores of the more critical accessions and those that have not been planted recently at the CWN. Stricter planting priorities created several years of backlogs in normally scheduled seed increases which led to more accessions being flagged as unavailable.

The characterization of the NCGC has been efficient because it is about 80% completed. Gaps in capturing ratings and images of mature plants remain in critical and photoperiodic accessions which makes the CWN essential for revitalizing this underutilized part of the NCGC as well as completing the characterization. The NCGC credits CottonGen for quick uploading of descriptors and images because it is a different process in GRIN-Global. The more data that is online, hopefully the more conservative are the accessions requested from the NCGC because demand for limited seed remains high.

Acquisition of new germplasm still proceeds cautiously because of the need to test for adventitious presence of transgenes, or of issues in getting permission to collect or exchange germplasm. A thorough knowledge of gaps in the NCGC better pinpoints germplasm that will add genetic diversity without risk of duplication.

The NCGC strives to meet the demands from many users, ranging from cotton breeders and scientists to home growers and educators. Collaboration with users helped us to complete characterization tasks, better address the germplasm needs of the scientific community, make contributions to research (Zhang et al. 2020, Huang et al 2020, Yu et al 2020), but also to get feedback from evaluations that better reveal the diversity and utility of the NCGC and hopefully better matches the right germplasm to the needs of the users.

References

Zhang, J., Abdelraheem, A., Zhu, Y., Wheeler, T.A., Dever, J., Frelichowski, J.E., Love, J., Ulloa, M., Jenkins, J.N., McCarty Jr, J.C., Nichols, R., Wedegaertner, T. 2020. Assessing genetic variation for Fusarium wilt race 4 resistance in tetraploid cotton by screening over three thousand germplasm lines under greenhouse or controlled conditions. Euphytica. 216:108. https://doi.org/10.1007/s10681-020-02646-2.

Huang, G., Wu, Z., Percy, R.G., Bai, M., Li, Y., Frelichowski, J.E., Hu, J., Wang, K., Yu, J., Zhu, Y. 2020. Genome sequence of *Gossypium herbaceum* and genome updates of *Gossypium arboreum* and *Gossypium hirsutum* provide insights into cotton A-genome evolution. Nature Genetics. https://doi.org/10.1038/s41588-020-0607-4.

Yu, J., Frelichowski, J.E., Hinze, L.L., Udall, J.A. 2020. Exploiting genetic variation of *Gossypium* gene pools for cotton improvement [abstract]. Plant and Animal Genome Conference, January 11-15, 2020, San Diego, California. W820.