Work Completed

- Curated marker data (increased from 129,937 markers to 338817, an increase of 160% in CottonGen) The curated marker data added includes marker name/species/alleles/sequence, such as primers, allele specific probes, variety, source of description (cDNA or genomic DNA), etc. Specifically this includes:
  - 208,880 new SNP and InDel markers for six species (G. hirsutum, G. barbadense, G. armourianum, G. mustelinum, G. longicalyx, G. mustelinum) developed by five research institutes (TAMU, UCD, BYU, CCRI, DOW):
    - 14,093 new G. hirsutum and G. barbadense markers from CCRI (13,967 SNPs and 126 InDels)
    - 32,113 G. hirsutum SNPs from UCD, DOW, and TAMU
    - 35,484 G. longicalyx specific SNPs from TAMU
    - 171 G. hirsutum SNPs with primers from TAMU
    - 34,550 G. longicalyx SNPs from TAMU
    - 6,396 G. tomentosum SNPs from TAMU
    - 6,669 G. mustelinum SNPs from TAMU
    - 24,829 G. armourianum SNPs from TAMU
    - 14,613 G. hirsutum SNPs from UCD
    - 39,962 BYU SNPs with new information from NCBI
  - 102,547 existing Cottongen SNPs were updated and reloaded:
    - 1052 BYU old SNPs
    - 35,265 CIR SNPs
    - 66,230 NBRI SNPs

- Map Data
  - 40 genetic maps extracted from 40 publications curated and uploaded to the comparative mapping viewer CMap, bringing the total to 92, an increase of 77%. Map, marker and QTL data associated with these maps ready to upload to the marker database in CottonGen. Map/marker and QTL data from a further 60 publications being collated for upload in Q4. So in Q3 and Q4, map, marker and qtl data from 100 publications will be added to CottonGen, making it current for these data types.

- Transcriptome Data RefTrans Analysis
  Individual reference transcriptomes under development for G. arboreum, G. raimondii, G. hirsutum and G. barbadense being developed using public RNASeq and EST data. Project 75% complete, will be finished and available in Q4 in time for Beltwide and PAG presentations.

- Breeding Information Management System
  Work ongoing to get example cotton data into BIMS and develop v1 for demo purpose in Q4.
➢ Other
  o VCF files from BYU Udall's Lab made available online available and data can be downloadable

➢ Usage in Q3 of 2016: 6,228 visits by 2,401 unique viewers from 100 countries who viewed 35,323 pages.

Work in Progress and plan in Q4, 2016

  o Completing v1 of the species refTrans
  o Adding QTL data for 100 publications.
  o Add genotype data for 400 lines for the 63K array
  o V1 of BIMS available to demo
  o Developing further terms for Trait Ontology to describe cotton traits.
  o Develop SNP genotype search interface