

High-Biomass Energy Crops for U.S. Energy Security

Submitted to:

DARPA

Submitted by:

William Rooney (PI), Professor, Soil and Crop Sciences
Bob Avant, Director, Bioenergy Programs
Jürg Blumenthal, Associate Professor, Soil and Crop Sciences
Nael El-Hout, Research Scientist, Texas AgriLife Research and Extension Center, Weslaco
Mike Gould, Director, Texas AgriLife Research and Extension Center, Weslaco
Russell Jessup, Assistant Professor, Soil and Crop Sciences
Patricia Klein, Associate Professor, Horticulture
John Mullet, Professor, Biochemistry and Biophysics
James Richardson, Professor, Agricultural Economics
Steve Searcy, Professor, Biological and Agricultural Engineering
David Stelly, Professor, Soil and Crop Sciences

Texas AgriLife Research
Texas A&M University System

Ceres, Inc.

Contacts:

Brett Giroir, Vice Chancellor for Research
Texas A&M University System

Bill F. McCutchen, Associate Director
Texas AgriLife Research
Texas A&M University System
bmccutchen@tamu.edu

October 16, 2009

High-Biomass Energy Crops for U.S. Energy Security

VISION: The Texas A&M University System (TAMUS) will create and deliver advanced high yielding energy sorghums and new energy crops through a novel, non-GMO, wide-hybridization technology platform. These unique energy crops will be sustainable, high yielding, widely adapted, drought tolerant, and optimized for biofuels and biopower generation, thereby significantly improving U.S. biofuels and biopower production capability and long-term energy security.

REQUIRED RESOURCES: \$21M / 5 years - - By year: \$3.3M, \$3.4M, \$5M, \$5M, \$5M.

PROJECT DELIVERABLES:

1. Energy sorghum production maximized in locations of national security importance using optimized management practices, harvest logistics, and economic assessment.
2. Energy sorghum hybrids with increased yield and optimized composition for advanced biofuels and biopower generation designed using an integrated genomics-to-energy crop breeding technology platform.
3. Novel wide-hybrid energy crops propagated vegetatively and/or through seed production and determination of the genetic basis of wide hybridization.

OVERALL APPROACH: The capacity to sustainably produce a large supply of low-cost compositionally optimized biomass is an essential requirement of large-scale economic production of advanced biofuels. Researchers at TAMUS have determined that energy sorghum and wide-hybrids derived by crossing sorghum to cane and other energy grasses can meet the biomass feedstock requirements of the cellulosic-advanced biofuels industry. The proposed R&D plan will accelerate the genetic improvement of energy sorghum, develop novel sorghum/cane wide-hybrids for energy use, deploy these crops at sites of national security importance, and optimize crop management, biomass harvesting and delivery to biorefineries.

NATIONAL SECURITY IMPLICATIONS: Energy sorghum hybrids optimized for geographic regions of national security importance have the capacity to supply lower cost feedstocks with specialized characteristics for advanced biofuels or biopower generation with high GHG offsets. Accelerating the improvement of energy crops will lead to cost decreases, supply increases, better conversion efficiency, reduced competition with food and feed production, and improved life-cycle assessment of fuels derived from energy sorghum and wide-hybrids. Successful implementation of these technologies will provide a strategic, sustainable, low-cost source of advanced biofuels and biopower for DOD, and in the longer term could significantly impact the world's supply of energy and transportation fuels.

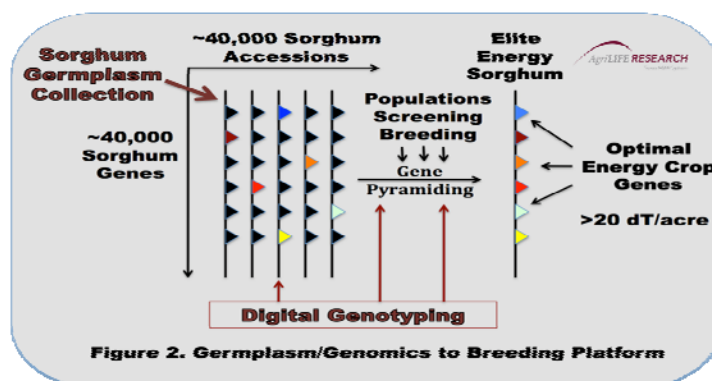
GOAL 1. *Grow and optimize production of energy sorghum hybrids in strategic areas to provide feedstocks for biofuels production for military installations (Figure 1).* TAMUS has created unique high-tonnage drought-tolerant, non-GMO “energy sorghum” hybrids that yield 10-15 dry tons (dT) of lignocellulosic biomass per acre. Current lines of energy sorghum enable

biomass delivery to biorefineries at ~\$60/dT providing a ~75% GHG offset for advanced biofuels production or ~95% for biopower. Energy sorghum hybrids will be grown at sites of national security interest such as Hawaii and other locations where agronomic management and harvest logistics will be optimized. Hybrid seed will be produced and tested in cooperation with our corporate partner Ceres, a commercial energy crop seed company.



Figure 1: Bioenergy Sorghum Production

GOAL 2. Increase the yield and optimize the composition of energy sorghum using an integrated genomics-to-breeding technology platform to accelerate the rate of genetic improvement and design of energy sorghum (Figure 2). Sorghum is one of the most highly productive C4 grass species known, with excellent genetics, a complete genome sequence, and an extensive germplasm collection (~40,000 accessions) that contains valuable genetic diversity for the design and improvement of energy sorghum and other energy crops. TAMUS is creating an advanced Genomics to Breeding Technology Platform based in part on a novel digital genotyping technology that enables the acquisition of detailed profiles of genetic variation of germplasm accessions, gene discovery populations, and breeding materials. Implementation of this technology within an expanded TAMUS energy sorghum and wide-hybrid genomics, breeding, and improvement program will significantly increase the long-term rate of energy crop design and improvement. The goal is to increase energy crop yield to 15-20 dT/acre while optimizing biomass composition to improve conversion efficiency into advanced biofuels and biopower.



GOAL 3. Create revolutionary next-generation energy crops by wide hybridization of diverse sorghums with sugarcane and other energy grasses (Figure 3). Wide-hybridization technology developed at TAMUS allows sorghum to be crossed with sugarcane and other C4 grasses. The resulting intergeneric hybrids have the potential for increased yield, unusual hybrid vigor, and unique combinations of useful traits from the species being crossed, such as large seed size or drought tolerance from sorghum, with perennial growth, cold tolerance, and/or high-nitrogen use efficiency from other C4 grasses such as energy cane or Miscanthus. Revolutionary ramifications are expected: [1] mass-production of hybrid seed from crosses with energy canes (first time ever); [2] 25-90+% cost reduction for planting energy canes; [3] option to grow wide-hybrid energy crops as seed-based crops (*much* more flexible than production based on vegetative propagules); and [4] a suite of newly created next-generation energy grasses for advanced biofuels and biopower generation in diverse agricultural and climatic conditions.

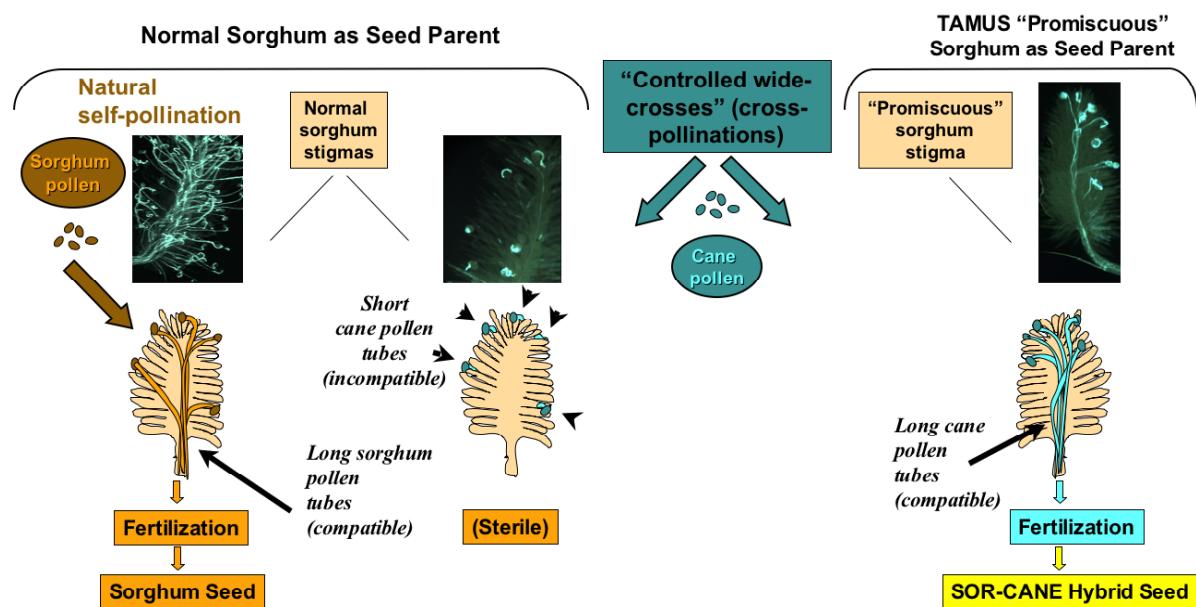


FIGURE 3. Diagrammatic representation of wide-hybridization technology for sorghum (*Sorghum bicolor*): normal sorghum stigmas are highly “self-compatible”, i.e., with sorghum pollen (left), but highly incompatible with alien pollen (center), pollen tubes of which are forced to cease growth soon after emergence, thus precluding fertilization. In the specially bred TAMUS “promiscuous” lines of sorghum, the alien pollen is much more compatible (right) resulting in pollen tube growth, which thus permits fertilization, embryogenesis and formation of F1 seed

IP/LICENSING PLAN: The intellectual property for sorghum wide-hybridization and wide-hybrids is wholly owned by TAMUS. Current energy sorghum hybrids (not wide-hybrids) are being developed in collaboration with Ceres, a commercial energy crop seed company. Ceres has agreed to negotiate preferred customer status with DOD for its sorghum hybrids and to supply commercial amounts of seed. The results of research funded by DARPA will be protected via both statutory means (e.g. patents, plant variety protection certificate) and contractual means (e.g. distribution of tangible materials under appropriate agreements). Licensing of TAMUS

intellectual property will include diligence to ensure licensees will meet both commercial and national security requirements for these novel energy crops.

RESEARCH AND DEVELOPMENT PLAN:

GOAL 1: Optimize production of energy sorghum hybrids in strategic areas to provide feedstocks for biofuels production for military installations.

Deliverables:

1. Energy sorghum hybrids adapted to specific production environments of national security interest.
2. Agronomic guidelines for sustainable production of energy sorghum.
3. Location-specific logistics system(s) for supplying biomass feedstocks to biorefineries that meet product specifications.
4. Economic assessment of dedicated energy sorghum production and biomass delivery to biorefineries at sites of national security interest.

Approach:

Energy sorghum, a C4 grass, is a superior plant species for lignocellulosic bioenergy production due to its ability to produce at least 10-15 dry tons/acre of biomass in good production environments. The development of high-yielding energy crops such as energy sorghum is important to the successful establishment of large-scale biofuel production because these crops will (1) reduce the cost of biomass per unit of biofuels produced, (2) reduce competition for land and between food and fuel crops by minimizing bioenergy production footprints, (3) reduce the cost of transporting biomass to biorefineries by increasing the amount of biomass that can be generated near a biorefinery, and (4) reduce carbon emissions per unit of transportation fuel produced, generating carbon credits and positive environmental impact.

The goal of this objective is to maximize the yield of energy sorghum hybrids in sites of national security interest by screening current and next-generation energy sorghum hybrids for adapting to specific locations, optimizing crop management, establishing harvest schedules, and developing the logistics of processing and transport of biomass feedstock to biorefineries. The economics of the integrated energy sorghum production system will also be analyzed.

Task 1.1: Identify the best energy sorghum hybrids for specific production environments of national security interest.

Rationale:

Energy sorghum hybrids need to be evaluated in multiple environments to identify the adaptation range of hybrids across the US, to identify the specific hybrids that are best adapted to key production sites, to assess the yield potential in each region or site, and to identify the factors that limit yield in each environment so they can be minimized through genetic improvement and crop management.

Work Plan:

1.1.1 Small replicated trials of at least 25 different energy sorghum hybrids in 8-10 environments to assess and identify the optimum genotypes for each adaptation zone. Production test sites will include and focus on sites of current national security interest such as Hawaii and West Texas and broader regions to maintain flexibility for planting in other locations that are likely to be important in the longer term. As appropriate to environment, selected energycane varieties will be included for relative comparison of performance.

1.1.2 Plant breeding data include but is not limited to:

- Biomass Yield
- Composition
- Height, Maturity
- Disease Reaction
- Pest Reaction

Cost 1.1:

<u>Year 1</u>	<u>Year 2</u>	<u>Year 3</u>	<u>Year 4</u>	<u>Year 5</u>	<u>Total</u>
300,000	300,000	350,000	350,000	350,000	1,650,000

Task 1.2: Establish sustainable and best management practices for optimum agronomic production of energy sorghum at selected locations of strategic importance in Texas and Hawaii.

Rationale:

The genetic potential of elite hybrids or cultivars will be expressed and yield maximized only if best management practices are utilized to optimize the production environment.

Work Plan:

1.2.1 Large replicated trials (10 – 20 acres) of the highest yielding energy sorghum hybrids for agronomic testing at selected locations. Variable factors to evaluate include:

- Plant population
- Fertility, irrigation requirements
- Hybrid or cultivar
- Composition

1.2.2 Long-term, large scale cropping systems experiments will be conducted to evaluate the effect of agronomic practices on cropping system performance and sustainability.

1.2.3 As appropriate to locations such as Hawaii, energycane cultivars will be planted and evaluated in comparable studies.

Cost 1.2:

<u>Year 1</u>	<u>Year 2</u>	<u>Year 3</u>	<u>Year 4</u>	<u>Year 5</u>	<u>Total</u>
216,009	218,363	447,512	452,579	458,030	1,792,493

Task 1.3: Enhance logistics systems to meet biomass feedstock specifications (quantities, moisture content, particle size, total carbohydrate composition, ash content, etc.) with acceptable cost and maximum energy gain.

Rationale:

Unit operations developed for other industries can deliver herbaceous biomass for energy conversion; however, those operations have high dollar and energy costs. New or redesigned unit operations are required to produce cost-competitive biofuels locally. Machines and processes developed as unit operations for other agricultural operations (animal feed, sugar, cotton, etc.) can deliver herbaceous biomass for energy conversion. However, those operations will have high operational and energy costs. Research in this project is focused on development and optimization of a feedstock logistics system for high yielding crops (energy sorghum and subsequently developed hybrids). Biomass logistics literature reports estimates of cost and energy consumption for the best current biomass logistics technology (based on large square bales) as a cost to collect and process of approximately \$65/dt and an energy gain ratio (energy content delivered/energy consumed in collection and processing) of 20. These values are for the feedstock supply chain only and do not include costs of producing or converting the biomass. In addition, these estimates are for crop residues and perennial grasses. No estimates are available for high yielding crops such as energy sorghum. These estimated values are a baseline that will be used to evaluate advances gained during the research.

Given DOD's desire for drop-in biofuels that can displace significant quantities of petroleum-based JP-8 and JP-5, the most likely conversion techniques are expected to be thermo-chemical in nature. Without selecting a specific conversion technique, general feedstock characteristics can be anticipated. Thermo-chemical conversion techniques typically require biomass with lower moisture content and smaller, more uniform particle sizes, both of which are significant challenges without applying excessively high levels of energy. All technical challenges for providing the desired feedstock characteristics will be prioritized with the greatest emphasis placed on the most limiting factors.

Work Plan:

- 1.3.1 Identify and evaluate logistics systems optimized for the two DOD specified biofuels production sites and the specified biomass characteristics required by conversion processes to produce JP-8.
 - a. Establish sites for producing energy sorghum. It is anticipated that one of these sites will be in Hawaii and a second will be near College Station, Texas, to represent a continental U.S. site.
 - b. Establish systems using best current technology appropriate for the selected sites.
 - c. Determine the probability of working days (days when weather allows field work to occur) for the selected sites using historical weather records.
 - d. Conduct field evaluations of the identified logistics systems.
- 1.3.2 Develop the ability to accurately simulate the performance of an energy sorghum logistics system using the Department of Energy's IBSAL modeling software.
 - a. Modify IBSAL to incorporate energy sorghum as a candidate biomass crop.

- b. Modify and validate functional relationships in the unit operation elements for accurate prediction with energy sorghum.
- c. Generate new or modify existing IBSAL elements to describe the unit operations developed under subtask 1.3.3.

1.3.3 Improve unit operations and interactions in the site-specific logistics systems to reduce the cost of delivered feedstock and increase the energy gain.

- a. Prioritize best current technologies for potential to reduce costs and increase energy gain. While the specifics of each location will determine the operations with the greatest potential for improvement, the most likely opportunities are reducing moisture content prior to packaging, maximizing dry matter capture, and packaging to facilitate efficient transport and storage.
- b. Modify machines that perform the high-priority unit operations in collaboration with those industry partners best able to implement those modifications at an appropriate scale. Where possible, unit operations will be refined by modifying existing platforms, making collaboration with manufacturers the most efficient means of rapid advances.
- c. Perform field evaluations of the modified unit operations and refine those operations through continuous product improvement techniques.

Cost 1.3:

<u>Year 1</u>	<u>Year 2</u>	<u>Year 3</u>	<u>Year 4</u>	<u>Year 5</u>	<u>Total</u>
252,948	416,206	855,241	850,987	740,733	3,116,115

Task 1.4: Determine the optimal production harvesting system in alternative regions and prices farmers must be paid to grow a dedicated energy crop.

Rationale:

Critical to the project's success is knowing what the cost of producing biomass feedstocks and the prices (and contracts) farmers will have to be offered to encourage them to grow a dedicated energy crop. Risk of production and harvesting energy crops must be considered relative to risks of producing traditional crops or breaking out new land to grow energy crops.

Work Plan:

The results of the field data collection and logistics system simulation will be combined into an assessment of sorghum production, harvest, and delivery system. This effort will include details on the equipment, fuel use, labor requirements, and costs of planting, growing, harvesting, handling, and transporting feedstock in order to generate accurate projections of delivered feedstock cost under a commercial scenario. Information about weather risk, soil productivity, and costs of production for traditional crops at each alternative site will be modeled in a Monte Carlo economic model to project risk-adjusted net incomes for energy crops and existing food/feed crops. Costs of production and production functions for dedicated energy crops will be modeled to estimate risk-adjusted net returns for energy crops using alternative production and harvesting practices. Stochastic efficiency techniques will use the risky net returns for energy crops and traditional crops to estimate utility-adjusted net returns for alternative price/contract arrangements to estimate which price/contract will be necessary to induce farmers

to grow energy crops. Model results will be validated through interviews with local commercial-scale farmers.

Analyses of acreage competition with traditional food/feed crops will be done for alternative areas. For regions with no current crop production, the costs to convert open ground to energy crop production will be estimated. Budgets developed for growing an energy crop under alternative production and harvesting systems will be used to determine the economically optimal production/harvesting system for each study area. Local risk factors are expected to make the optimal production/harvesting system differ across study areas.

Cost 1.4:

<u>Year 1</u>	<u>Year 2</u>	<u>Year 3</u>	<u>Year 4</u>	<u>Year 5</u>	<u>Total</u>
224,999	225,000	341,785	350,000	350,000	1,491,784

Task 1.5: Identify feasible production regions for commercial production of dedicated energy crops at sites of national security interest.

Rationale:

Many different factors must be considered in identifying production sites. Site selection is a critical factor, as it will directly affect what kind of energy sorghum to produce, crop yields, and the costs of production. Early selection of feasible sites will directly affect the plant breeding program and locations of field tests. Once feasible sites are identified, the basic and agronomic research can proceed to optimize the energy crop's biomass production potential for regions of interest. The site selection criteria must consider many different factors, such as growing season for sorghum, current production of sorghum in the region, harvesting and storage requirements, proximity to end users, availability of land and irrigation water, rainfall patterns, temperature patterns, competition with existing crops and livestock, population density, and environmental considerations.

Work Plan:

1.5.1 Identify the DARPA site selection criteria.

1.5.2 Identify additional location criteria based on secondary data and experience and assign weights to site selection criteria to form a lexicographic function for ranking alternative sites.

1.5.3 Develop models to analyze the possible sites discussed at the meeting plus additional sites that also may meet the site selection criteria.

1.5.4 Location studies will be updated in subsequent years as the breeding program produces new varieties that could expand the climatic/resource zones where dedicated energy sorghum can be produced.

Cost 1.5:

<u>Year 1</u>	<u>Year 2</u>	<u>Year 3</u>	<u>Year 4</u>	<u>Year 5</u>	<u>Total</u>
75,000	75,000	50,000	50,000	50,000	300,000

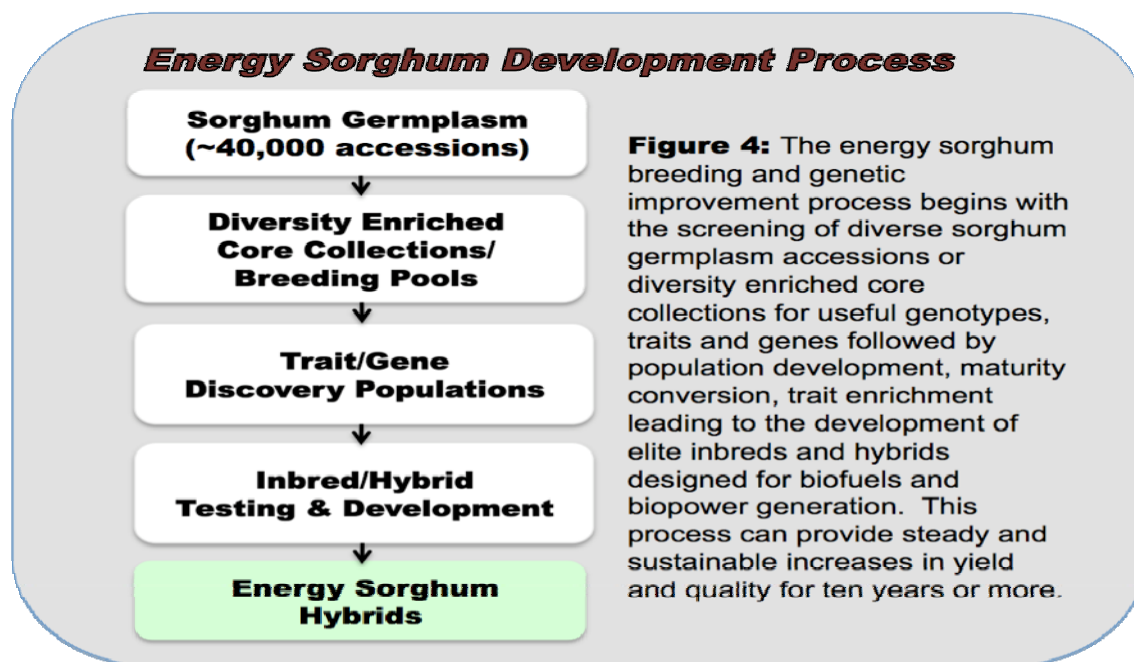
GOAL 2: Increase the yield and optimize the composition of energy sorghum using an integrated genomics-to-breeding technology platform to accelerate the rate of genetic improvement and design of energy sorghum.

Deliverables:

1. Energy sorghum inbreds and hybrids with increased yield (~15-20 dT/acre) and improved composition for DOD's biofuels/biopower conversion processes.
2. A genomics-to-breeding technology platform that will sustain increased rates of energy sorghum improvement.
3. Identification of genes/markers that control traits important for the design of energy sorghum and useful for marker-assisted improvement of energy crops.

Rationale and Approach: Substantial improvement in energy crop yield and quality can be achieved using breeding and genetic improvement methods that have been used to improve crops such as corn (~4X) over the past 50 years. The central approach involves screening germplasm collections for variation in genes/traits that contribute to increased yield and improvements in composition and performance (Figure 4). The breeding-based genetic improvement process is time consuming and labor intensive because of the complexity of germplasm collections, the structure and nature of naturally occurring genetic variation, and due to the need to simultaneously improve many different traits. However, the resulting improved genotypes are stable and provide sustainable increases in yield and quality.

The TAMUS sorghum-breeding program is developing next-generation energy sorghums using the development process shown in Figure 4, by screening the sorghum germplasm collection for useful genes/traits and then pyramiding these genes and traits through breeding and marker-assisted selection into elite energy sorghum hybrids. A key rate limiting step in energy sorghum improvement is the time and cost of screening the germplasm collection of ~40,000 accessions, core collections, and breeding populations for the large number of traits and genes that could improve energy crop design.



Analysis of the energy sorghum development and design process showed that the overall efficiency of genetic improvement could be significantly improved through the systematic collection and utilization of genotyping information. For example, the collection of genotyping information on a significant portion of the germplasm would eliminate the screening of duplicate accessions, enable the construction of diversity enriched core collections and breeding pools, aid trait locus discovery and DNA marker development, and enable marker-assisted breeding. To address this need, TAMUS developed Digital Genotyping Technology taking advantage of the capacity of new high performance sequencers such as the Illumina GAII to acquire billions of base pairs of DNA sequence information each run. A bioinformatics DNA sequence processing program was developed that can reveal the nature and distribution of genetic diversity at the DNA sequence level across the genomes of large numbers of accessions and breeding lines at variable depths depending on application. We propose to use this technology to assess the genetic diversity of ~16,000 germplasm accessions and breeding lines, create core collections enriched in genetic diversity, identify parental lines for population construction, for QTL/trait/gene mapping, and to carry out genome wide marker-assisted-breeding (MAB) (Figure 5). Implementation of Digital Genotyping within the integrated energy sorghum improvement program shown in Figure 5 will increase the rate and efficiency of energy sorghum development for at least the next ten years that we estimate will be required to fully utilize the genetic potential of the sorghum germplasm collection for energy crop design.

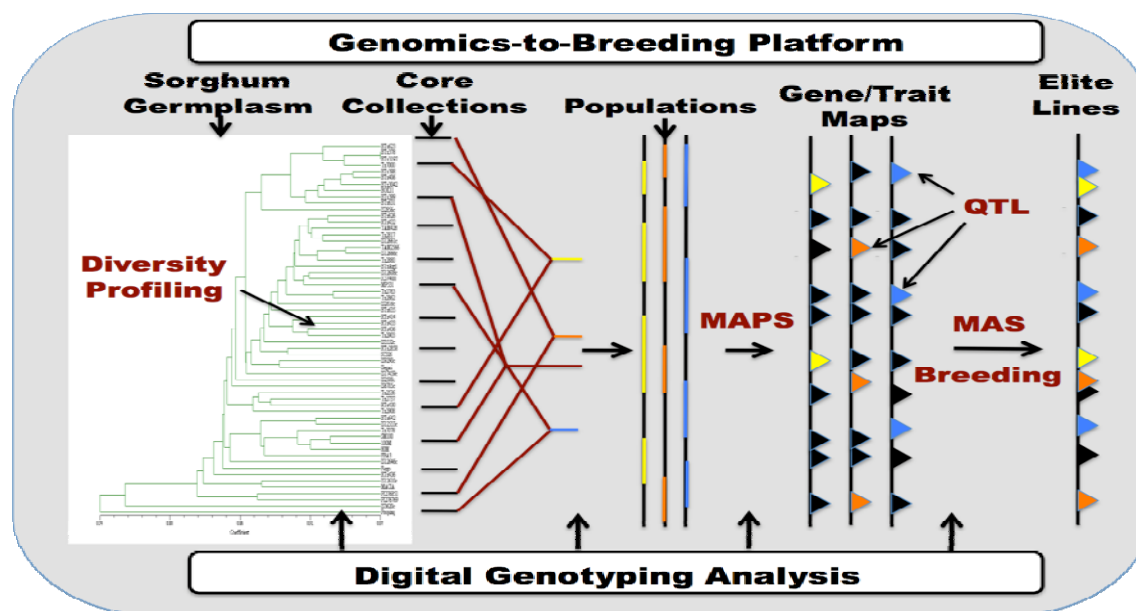


Figure 5: Genomics-to-Breeding Technology Platform that begins by collecting diversity profiles of sorghum germplasm accessions, followed by the creation of core collections, populations, gene/trait (QTL) mapping, marker-assisted selection (MAS) and breeding.

RESEARCH PLAN:

Task 2.1: Collect digital genotyping diversity profiles of ~16,000 sorghum germplasm accessions to enable more efficient and comprehensive utilization of germplasm for trait/gene discovery and energy sorghum breeding.

Work Plan:

2.1.1. Collect and analyze genomic diversity profiles of ~2,000 to ~4,000 sorghum germplasm accessions each year.

Digital Genotyping technology (DG) developed at TAMUS is based on the targeted acquisition of informative sequences from plant genomes of different accessions or lines, and the subsequent comparison of these sequences to identify DNA polymorphisms that reveal genetic relationships. In this project, DNA will be extracted from 2,000 (Yrs 1,2) or 4,000 (Yrs 3-5) sorghum accessions each year for genotyping and to build a DNA archive for subsequent use. Approximately once per month, DG templates will be prepared from the DNA derived from 200 (Yrs 1,2) or 400 (Yrs 3-5) sorghum accessions and sequenced using an Illumina GAI sequencer. Approximately 40bp of sequence will be acquired per template and each set of templates will be sequenced at ~10X depth. DG-1 depth of analysis will be carried out where approximately 15,000 unique sequences from specific sites in each sorghum genome will be analyzed revealing the distribution of up to ~2,000 DNA polymorphisms among diverse accessions. The sequence data will be processed using a bioinformatics DNA sequence analysis pipeline that removes non-specific sequence reads, clusters the sequences, and identifies sequences that are unique in the genome and therefore useful for genetic analysis. Data from the unique sequence sites will be placed in a database for ongoing analysis of genetic diversity. Once every 12 months, data collected on accessions will be analyzed for overall genetic relationships among germplasm accessions using standard programs for hierarchical clustering and visualization. This

information will guide the selection of sorghum accessions placed in core collections and analyzed at greater depth (DG-2) to reveal haplotypes (Objective 2.1.2).

2.1.2. Collect 200-400 high-resolution genotype profiles of sorghum accessions that comprise core collections, breeding lines and parental lines used in mapping each year.

DG-2 depth genotyping profiles will be obtained from sorghum genotypes selected to incrementally sample germplasm collection diversity, accessions selected for inclusion in core collections, discovery populations, and breeding lines. Approximately 200 accessions will be analyzed at DG-2 depth in Years 1 and 2 and 400 accessions/yr in Years 3-5. DG-2 depth of analysis involves the collection of sequences from ~120,000 sites that will identify more than 15,000 polymorphisms that distinguish the genotypes. This depth of genotyping will enable high-resolution graphical genotypes to be constructed revealing genomic regions that may contain allelic variation in key genes for energy sorghum traits. It should be noted that the numbers of genotypes projected for analysis in Years 1 and 2 are based on current Digital Genotyping and Illumina sequencing technology. However, we project that by Year 3, the number of genotypes that can be analyzed per sequencing run will increase due to improvements in the GAII sequencer and this increase in expected efficiency is reflected in the number of genotypes that we estimate can be collected during Years 3 to 5.

Cost 2.1:

<u>Year 1</u>	<u>Year 2</u>	<u>Year 3</u>	<u>Year 4</u>	<u>Year 5</u>	<u>Total</u>
250,000	251,708	346,199	349,806	352,288	1,550,001

Task 2.2: Create and screen core sorghum germplasm collections and an integrated set of breeding and discovery populations for phenotypes, genotypes, and genomic regions that contribute to yield, and QTL/genes that modulate key energy sorghum traits.

Work Plan:

2.2.1. Create and screen core sorghum germplasm collections for key energy traits including but not limited to biomass yield, composition/conversion efficiency, duration of growth, lodging resistance, and drought tolerance.

The goal of this activity is to capture a large portion of the genetic and trait diversity present in the sorghum germplasm collection in core collections suitable for screening energy traits, in breeding pools, and in discovery populations in order to accelerate comprehensive germplasm screening for genetic variation in key energy traits. An energy sorghum core collection will be developed, refined and enriched that contains a stratified nested core collection of photoperiod sensitive energy sorghum accessions (100, 400, 1,600 accessions having different degrees of relatedness) and a collection of converted (photoperiod insensitive) energy sorghum accessions that can be used to assess traits expressed in energy sorghum hybrids (100). The core collection will be screened for useful variation in a wide range of energy traits including but not limited to biomass yield, duration of growth, lodging resistance, composition/conversion efficiency, stem traits, and drought tolerance to identify parental lines useful for creating trait/marker discovery populations and genotypes useful for energy sorghum breeding (Task 2.3).

2.2.2. Create and screen energy sorghum populations for trait/gene locus mapping.

Populations segregating for a wide range of energy traits will be established to facilitate QTL/trait mapping and to delineate the complexity of the genetic architecture of the gene regulatory networks that modulate trait expression. Three populations consisting of ~96 lines derived from late flowering, photoperiod sensitive energy sorghum parental lines have been established and are at the F3 stage of construction. These populations will be advanced to the F4 (2010) and F5 (2011) generation. Preliminary screening for trait variation in these populations will be done in Yr 1 with deeper analysis in Yrs 2-5. Additional populations derived from crosses created during energy R-line conversion will be expanded and used for trait/QTL mapping. For example, a cross of [REDACTED] and a promising exotic R-line that is currently at the F2 stage will be expanded and progressed to subsequent generations. One set of progeny from this cross have a 'converted', early flowering genotype that allows test crossing to A-lines and evaluation as energy hybrids (late flowering). The converted lines will also be evaluated for traits per se. Another set of progeny derived from this cross are late flowering and these lines will be advanced to the F4/F5 stage so that energy sorghum traits can be evaluated by QTL mapping per se in a late flowering background. Progeny from both sub-populations will be genotyped using DG, and phenotyped for a range of traits including but not limited to biomass yield, flowering time, lodging resistance, stem traits, composition (NIR), and conversion efficiency (in DARPA preferred processes). The populations will be screened for additional traits such as drought tolerance, salt tolerance, nitrogen use efficiency and pest resistance as time and resources allow. Progeny from the [REDACTED] exotic R-line population will be evaluated row tests in College Station in 2011, and with replication per se and in hybrid combination in Yrs 3-5 assuming useful trait segregation is observed. Two additional populations with this genetic structure or derived from two converted energy sorghum R-lines that capture as much of the genetic diversity entering the breeding pool as possible will be generated each year then screened and advanced as described above. Therefore, within 5 years, at least 10 trait/marker discovery populations consisting of at least 96 progeny will have been created and will be at various stages of completion (F5 RILs).

2.2.3 Map QTL and generate markers for energy trait marker-assisted breeding.

Genotyping and phenotyping data obtained from the populations developed in objective 2.2.2 will be used to construct genetic and QTL maps using QTL Cartographer and standard mapping approaches. DNA markers for QTL that modulate useful energy traits identified through mapping will be generated starting in Year 3 and used for marker-assisted breeding once the markers are validated as useful in breeding pools (Task 2.3). QTL mapping will begin with one population in 2011 and scale up to two populations per year in Years 3-5. This information will provide the starting point for iterative searches for optimal alleles in the germplasm.

Cost 2.2:

<u>Year 1</u>	<u>Year 2</u>	<u>Year 3</u>	<u>Year 4</u>	<u>Year 5</u>	<u>Total</u>
249,477	249,539	396,091	400,456	404,437	1,700,000

Task 2.3: Develop and test advanced inbred lines and hybrids for yield, adaptation, and traits that contribute to the value and overall utility of energy crops.

Work Plan:

- 2.3.1 Construct a large number of crosses among converted energy lines to sample diversity and to utilize that diversity to enhance productivity of sorghum as a bioenergy crop. Population will be developed for both QTL mapping and breeding approaches.
- 2.3.2 Test-cross/phenotype/genotype/select improved lines. Selection and evaluation of these derived lines will occur over years and several locations. Traits that will be evaluated include, but are not limited to,
- a. Biomass yield/partitioning
 - a. Tillering, lodging
 - b. Flowering time, height
 - c. Composition (NIR)
 - d. Conversion efficiency
 - e. Drought/salt tolerance
- 2.3.3 Conduct selective mapping per se and in hybrids.
- 2.3.4 Pyramid beneficial traits in elite pollinator parents.
- a. Using phenotypic selection.
 - b. Using marker-assisted-selection.
- 2.3.5 Generate and test hybrids in multiple locations.

Cost 2.3:

<u>Year 1</u>	<u>Year 2</u>	<u>Year 3</u>	<u>Year 4</u>	<u>Year 5</u>	<u>Total</u>
527,152	491,963	775,956	723,208	731,724	3,205,003

GOAL 3: *Create revolutionary next-generation energy crops by wide hybridization of diverse sorghums with sugarcane and other energy grasses.*

Deliverables:

1. Elite intergeneric hybrids individually selected for vegetative propagation and biofuel production characteristics; diversity that expands production range.
2. Assessments of values and production ranges for the elite intergeneric hybrids.
3. Parental lines that optimize production of intergeneric F1 or BC1F1 seed for seed-based cultivation of heterotic “sor-canes” and other wide-hybrids, and/or self-fertile synthetic polyploids.
4. Genetic types that expand the ease and scope of wide-hybridization among C4 grasses for deriving additional biofuel crops.

Approach:

Because the highest yields of plant biomass are derived from exceptional C4 grass hybrids, *the most energy- and cost-effective plant-based biofuel strategies will rely on C4 hybrids.* TAMUS

recently established technology that revolutionizes the range and facility of wide-hybridization between sorghum and other C4 grasses, including sugarcane. TAMUS is *uniquely* capable of creating extremely large numbers of genetically unique intergeneric hybrids that have *never* been produced before. These have the potential for [1] strong vegetative hybrid vigor, [2] unique combinations of useful traits from the species and genotypes being crossed, [3] high capacities for clone-based production, like sugarcane, *and* [4] large enough seed sizes to allow for production of “wide-hybrid” canes from seed-based plantings. Potential trait combinations include resistances to multiple diseases, pests, drought, heat, cold, salinity, and soil deficiencies, perennial or annual growth, high nitrogen-use efficiency. This DARPA project will expectedly lead to a suite of novel next-generation energy grasses for advanced biofuels and biopower generation, the revolutionary benefits of which will include significantly higher biofuel yields, greater diversity of biofuel crops available to DARPA and other bio-fuel producers, expanded geographic, climatic and environmental ranges usable for biofuel production (thereby reducing impact on food production), and a seed-based system for producing canelike intergeneric crops which eliminate costly vegetative plantings (up to \$1,000/acre).

Task 3.1: Create large numbers of individually distinct sorghum/energycane hybrids.

Rationale:

Evaluation of current wide-hybrids indicates that “elite” hybrids (top 2%) offer great promise as biofuel feedstocks. If grown clonally as biofeedstocks using methods that rely on vegetative propagules, as for sugarcane, only the very best wide-hybrids will be used for biofuel production. The best hybrids will be obtained most rapidly by screening very large numbers of progeny from very diverse pollen parents. The best hybrids will be advanced for use in crop testing and hybrid development.

Work Plan

- 3.1.1 Hybridization of [REDACTED] sorghum (and other [REDACTED] derivative lines) with energycane and numerous, high-biomass grass species with varied regional adaptation in crossing greenhouses in Weslaco and/or College Station, Texas.
- 3.1.2 Produce at least 10,000 intergeneric seed annually
- 3.1.3 Selectively establish at least 1,000 seedlings in field nurseries in College Station for evaluation.
- 3.1.4 Over five years, screen at least 50,000 hybrids. Select hybrids will be propagated at the end of the season for use in other tasks.

Cost 3.1

<u>Year 1</u>	<u>Year 2</u>	<u>Year 3</u>	<u>Year 4</u>	<u>Year 5</u>	<u>Total</u>
176,202	150,668	208,889	204,130	204,158	944,047

Task 3.2: Multi-location screening of wide-hybrids to identify elite and unique hybrids.

Rationale:

The performance of intergeneric hybrids will be affected by interactions of genotype and environment (GxE). Given the possibility that subsequent production may be best served by selection directly in each area of subsequent use, we will grow and select early-generation wide-hybrids in multiple locations that are expected to be biofuel production sites. Data and selections from these sites will collectively reveal site-specific and site-nonspecific trait and/or genetic tendencies that foster improved selection subsequently.

Work Plan:

3.2.1 Intergeneric hybrids will be grown as transplanted seedlings or direct-seeded seedlings in three different Texas production environments (College Station, Weslaco, and Beaumont). One of two general approaches will be used to distribute materials to multiple locations destined to serve as biofuel production test sites.

3.2.2 F₁ seedlings selected in breeding nurseries will be vegetatively propagated and distributed clonally for replicated evaluation at the test sites.

3.2.3 Wide-hybrid F₁ seed identified as to parentage will be distributed to the multiple-location sites, for either direct seeding or greenhouse-based germination and subsequent transplanting; in this case, results from multiple locations would be collated according to parentage, not individual genotypes (i.e., family-based selection).

Cost 3.2:

<u>Year 1</u>	<u>Year 2</u>	<u>Year 3</u>	<u>Year 4</u>	<u>Year 5</u>	<u>Total</u>
150,049	149,996	199,909	199,468	200,599	900,022

Task 3.3: Determine agronomic performance and conduct pilot-scale testing of intergeneric hybrids.

Rationale:

Agronomic characterization of intergeneric hybrids will define the best management practices for these new crops, their relative strengths, and the best genotypes for commercialization. This information will be used to identify clones to advance for research Task 4.

Work Plan:

3.3.1 Years 1-2:

- a. Hybrids selected in Task 1 or 2 will be clonally propagated and planted in multiple locations (Weslaco, College Station, and Beaumont, Texas).
- b. In select locations, two complete trials will be planted; one will be grown under full irrigation and the second will be produced under limited irrigation to fully assess the level of drought tolerance.
- c. Evaluation for at least three years to determine adaptation, perenniality and biomass yield. All trials will have comparative checks of bioenergy sorghum and energycane.

3.3.2 Years 3-5:

- a. Continue additional testing, as described in Years 1-2.
- b. In years 3-5, the elite lines will be clonally propagated for advanced testing in a larger number of locations; including but not limited to sites of importance to DARPA and DOD production.

Cost 3.3:

<u>Year 1</u>	<u>Year 2</u>	<u>Year 3</u>	<u>Year 4</u>	<u>Year 5</u>	<u>Total</u>
149,986	220,055	300,187	301,139	302,147	\$1,273,514

Task 3.4: Develop seed-based propagation systems for sorghum/energycane hybrids.

Rationale:

Seed-based planting of an energycane-like crop could reduce planting costs by >90% and greatly increase production flexibility - - an attractive feature for both producer and biofuel plant feedstock production facilities. Current data indicate that *major* strides in seed production and hybrid performance will be obtainable through improvements of both parents, i.e., breeding of improved sorghum seed parents and concomitant selection among energycane pollen parents.

Work Plan:

3.4.1 Sorghum Seed Parental Line Development ()

- a. () sorghum with high stem sugar
- b. Enhanced disease resistance
- c. Improved drought tolerance
- d. Production of A/B lines

3.4.2 Energycane Pollinator Parent Screening and Development

- a. Assess seed set of different energycane pollinators on () or other () sorghums
- b. Select energycane pollinators based on pollen viability, () seed set, and progeny performance
- c. Selectively breed energycane specifically for pollination of () sorghum genotypes
- d. Test selected derivatives from breeding for enhanced pollination of () sorghum genotypes

3.4.3 Cytological Manipulations of Pollinator Parent to Produce Uniform Commercial Hybrids

- a. Recover an amphiploid of each elite sorghum/energycane hybrid
- b. Test fertility of derived line
- c. Test hybridization of derived line to sorghum (4.1a)
- d. Derive lines that are good pollinators and genetically stable

Cost 3.4

<u>Year 1</u>	<u>Year 2</u>	<u>Year 3</u>	<u>Year 4</u>	<u>Year 5</u>	<u>Total</u>
635,249	509,363	700,308	691,186	692,493	3,228,599

Task 3.5: Molecular characterization of [REDACTED] and other important genes that influence intergeneric hybridization and recombination, and parental and intergeneric hybrid genotypes.

Rationale:

Characterization of the [REDACTED] gene will facilitate further selection of promising parental genotypes for producing intergeneric crosses. In this case, marker-assisted selection will expedite and economize the development of genetic combinations of [REDACTED] with other genes that enhance production, performance, and quality of wide-hybrids. Mapping of a trait-controlling gene at high resolution is often key to its identification, which may well stimulate the discovery and utilization of similar genes in other species. Such advances could lead to additional advances in US biofuel feedstock production. We have developed a BC₁F₁ population consisting of ~300 progeny for initial mapping of the [REDACTED] locus. The BC₁F₁ population segregates for the *ms3* (male sterility) as well [REDACTED] to aid in phenotypic analysis. To begin map-based cloning of the [REDACTED] gene, the progeny from the BC₁F₁ population will be phenotyped for maize pollen tube growth using current cytological methods. Once phenotypic screening is completed, the parents and selected progeny (~12 [REDACTED] and 12 [REDACTED] *iap*), will be subjected to digital genotyping analysis at DG-1 depth to confirm the genomic location of the [REDACTED] locus. Simultaneously we will be examining additional maize pollen sources on sorghum [REDACTED] and [REDACTED] testers to develop a more rapid and efficient [REDACTED] screening system since currently the cytological method is extremely labor and time intensive. To continue fine mapping the [REDACTED] gene, an additional mapping population with ~1000 progeny will be constructed. Progeny from this high resolution will be phenotyped and genotyped to refine the [REDACTED] locus to less than 0.1cM (~3-4 genes). If necessary, additional markers flanking the [REDACTED] locus required for fine mapping will be developed from the sorghum pseudomolecule sequence. Once the locus has been delimited, potential candidate genes will be identified based on automated gene annotation, sorghum EST information, and comparative analysis with the syntenic regions of maize and rice. Candidate genes with the expected function will be sequenced (both genomic DNA and cDNA) to identify sequence polymorphisms that modify gene function. The sequences in the [REDACTED] loci will be examined for polymorphisms that could modify gene expression (mutation in promoter elements, altered RNA splicing/processing) or protein coding (amino acid change, protein function). To quantify gene expression at the RNA level, [REDACTED] candidate gene transcripts will be characterized by RT-PCR or mRNA sequence analysis on the Illumina GAII in the mapping population parents in vegetative tissues as well as in female reproductive tissues including stigma, style, and ovary. The combination of fine mapping, candidate gene allele sequence analysis, and expression profiling, will likely identify the [REDACTED] gene that is responsible for intergeneric hybridization in sorghum.

Digital genotyping profiles from selected cane genotypes that have been tested for efficacy in intergeneric hybridization will be performed to determine if this genotyping technology can be used to enhance parental line selection in the wide hybridization breeding program. DNA from selected cane genotypes will be extracted (yr 2) and analyzed at DG-1 depth. Approximately 40bp of sequence will be acquired per template, and on average each set of templates will be sequenced to 10X depth. Sequences will be processed through our DNA pipeline as previously

described. If initial digital genotyping analysis is successful with cane, full-scale implementation on ~400 intergeneric hybrids will begin in year 3 and continue for the duration of the project. DNA will be extracted from ~200 intergeneric hybrids twice a year and following DGA template preparation sequenced at DG-1 depth. Additionally, ~50 selected hybrids/year will be sequenced at DG-2 depth. Following DNA sequence processing, the unique sequences will be analyzed and archived in a database. This information will be used to determine the sizes and extents of genome admixture in the hybrids and will aid in the selection of intergeneric hybrids to move forward in the breeding program.

Work Plan:

3.5.1 Map and clone the [REDACTED] gene responsible for intergeneric hybridization in sorghum.

- a. Evaluate a BC₁F₁ mapping population (N=300) for the [REDACTED] phenotype.
- b. Genotype the BC₁F₁ mapping population to the ~0.3 cM resolution.
- c. Develop a rapid phenotypic screening method for [REDACTED] individuals.
- d. Construct a larger mapping population of ~1000 progeny to delimit the [REDACTED] locus to less than 0.1cM.
- e. Identify potential candidate genes within the [REDACTED] locus and sequence these genes in the parents of the [REDACTED] mapping population to identify functional differences.
- f. Validate the [REDACTED] candidate gene bioinformatically and experimentally as the gene that enables intergeneric hybridization.

3.5.2 Collect digital genotype profiles from selected cane genotypes as well as ~1200 intergeneric hybrids.

- a. Collect DNA from cane genotypes used in intergeneric hybridization as well as from ~400 intergeneric hybrids during years 3-5.
- b. Collect digital genotyping sequences at DG-1 depth.
- c. Process sequence data to identify SNPs and INDELs.
- d. Determine if DGA can enhance cane parental line selection and determine the sizes and extents of genome admixture in the intergeneric hybrids.
- e. Collect digital genotyping sequences at DG-2 depth from ~150 intergeneric hybrids to allow for visualization of haplotype variation among the hybrids.

Cost 3.5

<u>Year 1</u>	<u>Year 2</u>	<u>Year 3</u>	<u>Year 4</u>	<u>Year 5</u>	<u>Total</u>
149,839	149,885	199,601	199,784	200,417	899,526

BUDGET SUMMARY: Total Costs

Task	Year 1	Year 2	Year 3	Year 4	Year 5	Totals
1.1	\$300,000	\$300,000	\$350,000	\$350,000	\$350,000	\$1,650,000
1.2	\$216,009	\$218,363	\$447,512	\$452,579	\$458,030	\$1,792,493
1.3	\$252,948	\$416,206	\$855,241	\$850,987	\$740,733	\$3,116,115
1.4	\$224,999	\$225,000	\$341,785	\$350,000	\$350,000	\$1,491,784
1.5	<u>\$75,000</u>	<u>\$75,000</u>	<u>\$50,000</u>	<u>\$50,000</u>	<u>\$50,000</u>	<u>\$300,000</u>
	\$1,068,956	\$1,234,569	\$2,044,538	\$2,053,566	\$1,948,763	\$8,350,392
2.1	\$250,000	\$251,708	\$346,199	\$349,806	\$352,288	\$1,550,001
2.2	\$249,477	\$249,539	\$396,091	\$400,456	\$404,437	\$1,700,000
2.3	<u>\$527,152</u>	<u>\$491,963</u>	<u>\$775,956</u>	<u>\$723,208</u>	<u>\$731,724</u>	<u>\$3,250,003</u>
	\$1,026,629	\$993,210	\$1,518,246	\$1,473,470	\$1,488,449	\$6,500,004
3.1	\$176,202	\$150,668	\$208,889	\$204,130	\$204,148	\$944,037
3.2	\$150,050	\$149,996	\$199,909	\$199,468	\$200,599	\$900,022
3.3	\$149,986	\$220,055	\$300,187	\$301,139	\$302,147	\$1,273,514
3.4	\$635,249	\$509,363	\$700,308	\$691,186	\$692,493	\$3,228,599
3.5	<u>\$149,839</u>	<u>\$149,885</u>	<u>\$199,601</u>	<u>\$199,784</u>	<u>\$200,417</u>	<u>\$899,526</u>
	\$1,261,326	\$1,179,967	\$1,608,894	\$1,595,707	\$1,599,804	\$7,245,698
Total	\$3,356,911	\$3,407,746	\$5,171,678	\$5,122,743	\$5,037,016	\$22,096,094