Title: COMPOSITIONS AND METHODS FOR ENHANCING PLANT BREEDING

Abstract: Methods for using genetic marker genotype to improve the process of developing plant varieties with improved performance are provided. Methods for selecting essentially homozygous breeding lines useful for the production of hybrid progeny with an agronomically desirable trait are provided. The invention further relates to methods for producing such homozygous plants and to hybrid seeds obtained by crossing such homozygous parental plants and to plants grown from said seeds.
COMPOSITIONS AND METHODS FOR ENHANCING PLANT BREEDING

CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims the benefit of priority of U.S. Provisional Application Serial No. 61/952,556, filed March 13, 2014, the contents of which are herein incorporated by reference in their entirety.

FIELD OF THE INVENTION

The present disclosure relates to the field of agriculture, in particular to new plants and processes for obtaining them. More specifically, the disclosure relates to methods for improving the efficacy of a plant breeding program such as, for example, the process of domestication of wild plant populations. In some embodiments, the plant breeding program is aimed at altering phenotypic traits for which associations with genetic markers can be established. The invention further relates to a method for producing hybrid seeds and to hybrid seeds thus produced.

BACKGROUND OF THE INVENTION

In recent years, molecular breeding has demonstrated promise for improving the breeding process and thereby enhancing the rate of genetic gain. In molecular breeding, molecular markers provide a basis for selections of parental, progeny or tester breeding lines. This process may also be used in conjunction with phenotype-based selection. With the inclusion of genetic markers, breeding has advanced from selection for economically important traits in plants based on phenotypic records of an individual and its relatives to the application of molecular genetics to identify genomic regions that contain valuable genetic traits. Further, inclusion of genetic markers in breeding programs has accelerated the identification and accumulation of valuable traits into germplasm pools compared to that achieved based only on phenotypic data.

For molecular breeding to be effective, the differences in marker genotypes must be heritably associated to one or more phenotypic or performance traits. These associations are established by correlating the marker genotypes to lines or populations segregating for one or more traits. Genetic marker are often used to identify plants that contain a desired genotype at one or more loci, and that are expected to transfer the desired genotype, along with a desired phenotype for one or more traits, to their progeny. Markers that are highly correlated with a
phenotype are assumed to be genetically linked to the trait, thus the marker can then be used as a basis for selection decisions in lieu of phenotypically evaluating the trait per se. Markers that are not correlated will be inherited independently of the trait and are not useful for selections, but can be valuable in comparing similarities and/or measuring genetic distances among varieties and lines.

One significant problem with breeding strategies that rely on phenotypic selection is that most agricultural traits of interest are controlled by more than one genetic locus, each of which typically influences the given trait to a greater or lesser degree. It has been widely reported that the vast majority of economically important phenotypic traits in domesticated plants are so-called quantitative traits. Generally, the term "quantitative trait" has been used to describe a phenotype that exhibits continuous variability in expression and is the net result of multiple genetic loci presumably interacting with each other and/or with the environment.

One of the consequences of multi-factorial inheritance patterns is that it can be very difficult to map loci that contribute to the expression of such traits. However, recent years have seen tremendous advances in the application of marker-assisted breeding techniques, on both the development of markers and the association of markers with phenotypes, or quantitative trait loci (QTL) mapping. Examples of DNA markers are Restriction Fragment Length Polymorphisms (RFLP), Amplified Fragment Length Polymorphisms (AFLP), Simple Sequence Repeats (SSR), Single Nucleotide Polymorphisms (SNP), Insertion/Deletion Polymorphisms (Indels), Variable Number Tandem Repeats (VNTR), and Random Amplified Polymorphic DNA (RAPD), as well as many others known to those skilled in the art. The development of sets of polymorphic genetic markers (e.g., AFLPs, RAPDs, RFLPs, SNPs, SSRs, etc.) that span the genome has made it possible to investigate what molecular breeders often refer to as "quantitative trait loci" (QTL or QTLs), as well as their numbers, magnitudes, and distributions.

For example, various experimental approaches have been successfully developed to identify and analyze QTLs. One such approach involves crossing two inbred lines to produce F1 single cross hybrid progeny, selfing the F1 hybrid progeny to produce segregating F2 progeny, genotyping multiple marker loci, and evaluating one to several quantitative phenotypic traits among the segregating progeny. The QTLs are then identified on the basis of significant statistical associations between the genotypic values and the phenotypic variability among the segregating progeny. In this approach, the parental lines of the F1 generation are generally required to be inbred lines with known linkage phases. Therefore, such an approach
has been deployed primarily in various breeding programs of domesticated crop species, e.g., canola, maize, sorghum, soybean, etc.

However, in the case of undomesticated plant species such as *Jatropha* and many other energy-dedicated crops, perhaps one of the biggest challenges in designing an effective breeding program is the undomesticated nature of the species. Most of the *Jatropha* parental lines currently used in large-scale production of *Jatropha* hybrids are heterozygous at many loci and thus lack genetic uniformity. Therefore, a sexual cross between two genetically dissimilar parents typically results in a heterogeneous population of F1 hybrids, with each individual plant from the same cross exhibiting a unique allelic segregation event. Additionally, *Jatropha* breeders have observed not only genotypic segregation among F1 individuals, but also dramatic phenotypic differences. Further, the process of inbreeding of parent lines is often time consuming, especially since the phenology of *Jatropha* parental lines does not necessarily correlate well with yields of the F1 hybrid progeny. A similar conundrum is well documented in maize and other domesticated crops where hybridization is involved.

Until recently, *Jatropha* was considered a common crop but was not typically used for farming in large plantations until the advent of commercial biofuel production. Several characteristics of *Jatropha curcas* make it one of the ideal plants for biodiesel production. These characteristics include its non-food crop status; the ability to grow in difficult conditions including arid and otherwise non-arable areas, leaving prime areas available for food crop production; and relatively fast oil production in typically 1-2 years after planting.

The demand for oil and fats is expected to increase dramatically with the increase in world population. In the last few years, the non-edible oils market has developed largely due to the increasing demand for biofuels. Among the non-edible oils, *Jatropha* oil has been acknowledged for its physical and physicochemical properties, found suitable for use as biodiesel feedstock and industrial applications. For example, crushing *Jatropha* seeds produces *Jatropha* oil, which can be processed to produce a biodiesel that meets the restrictive specifications of EU biodiesel fuel standard (EN 14214), which reportedly can be used in a standard diesel car. Moreover, *Jatropha* oil is reported to fit well within the specification requirements of the main providers of endothermic engines, which generate electricity and heat to run straight on vegetable oil.

In addition, *Jatropha curcas* is gaining commercial importance as a biodiesel plant, and has attracted attention in terms of developing the plant in wastelands and dry lands. The increasing interest in oil from *Jatropha curcas* has generated enormous pressure to supply enough seeds that are homogenous and productive enough for plantation. As the need for
biofuels increases, the higher productivity of Jatropha makes it even more attractive as a target for biotechnological improvement. There is a need to increase the quality and yield of Jatropha oil and to rapidly develop useful characteristics when required, such as traits that enable the plant to grow in more arid locations, with higher oil yield and lower height increments.

However, although Jatropha is known for having wide adaptability and plethora of uses, the full potential of Jatropha is far from being realized because of several reasons. Relatively little research has been done to improve Jatropha varieties because the vast majority of Jatropha is produced by subsistent farmers. The crop's growth is therefore limited by insect pests, disease and weeds, rather than the plant's inherent ability. Several pests and diseases have been reported for Jatropha, e.g. the seed-feeding Scutelleridae, Agonosoma trilineatum; the scutellarid bug, Scutellera nobilis; and the inflorescence and capsule-borer, Pempelia morosalis. Typically, fruits of Jatropha in one bunch will ripe at different times, which cause difficulties during harvesting and leads to higher labor inputs. Therefore, new Jatropha varieties with uniform fruit maturation and ripening would be desirable and more suitable for large-scale mechanical harvesting methods. In recent years, research in Jatropha germplasm has increased, and a host of different Jatropha accessions from many different geographic areas are being assessed for desirable traits. However, very few of the desirable traits have been introgressed into commercial planting materials. In addition, given the previously limited knowledge about Jatropha genetics, a number of steps are required in order to improve germplasm domestication programs.

Thus, there is a long-standing and continuing need for new methods for optimizing breeding strategies for producing essentially uniform hybrid progeny with agronomically desirable genotypes. This and other needs are addressed by the presently disclosed subject matter.

**SUMMARY OF THE INVENTION**

The following embodiments and aspects thereof are described in conjunction with systems, tools, and methods which are meant to be exemplary and illustrative, not limiting in scope. In various embodiments, one or more of the above-described problems have been reduced or eliminated, while other embodiments are directed to other improvements. Any embodiment discussed herein with respect to one aspect of the invention applies to other aspects of the invention as well, unless specifically noted.

The present invention provides in a first aspect a method for selecting essentially homozygous parental plant lines from genetically-divergent, heterozygous plant lines for the production of hybrid progeny, comprising identifying one or more regions of the genomes of
said essentially homozygous parental plant lines that, when present together in said hybrid progeny, result in improvement of a target trait.

In one embodiment of the invention the method of identifying comprises the steps of crossing a pair of genetically-divergent, heterozygous plant lines to create an offspring population segregating for said target trait; selecting from said segregating offspring population of a first offspring subpopulation and a second offspring subpopulation, wherein said first offspring subpopulation comprises a plurality of positive transgressive segregants, and said second offspring subpopulation comprises a plurality of negative transgressive segregants; identifying in the genomes of said segregants one or more regions of said genomes that are essentially associated with said first offspring subpopulation, but that are essentially absent in said second offspring subpopulation and selecting one or more genetically-divergent, heterozygous plant lines comprising genomic regions that are identical or homologous to said one or more regions of said genomes identified; and breeding said one or more genetically divergent, heterozygous plant lines to produce essentially homozygous parental plant lines in which said homologous genomic regions have been fixed by said inbreeding.

In a second embodiment of the methods described herein, the inbreeding step of the selection method as described herein is performed by repeated selfings and/or or sib-crossings.

In a third embodiment of the methods described herein, the one or more regions of said genomes are genetically associated with at least one genetic molecular marker.

In a preferred embodiment of the methods described herein, the at least one genetic molecular marker is selected from the group consisting of a SNP marker, an AFLP marker, a DAF marker, a RAPD marker, an FLP marker, and an SSR marker.

Preferably, in another embodiment of the methods described herein of the essentially homozygous parental plant lines produced are homozygous for said one or more regions of genomes that, when present together in hybrid progeny, result in an improvement of said target trait.

In another preferred embodiment of the methods described herein, the genetically-divergent, heterozygous plant lines are monocotyledonous or dicotyledonous plant lines.

In another preferred embodiment of the methods described herein, the genetically-divergent, heterozygous plant lines are lines of a domesticated or undomesticated crop.

In yet another preferred embodiment of the methods described herein, the genetically-divergent, heterozygous plant lines are Jatropha lines, maize lines, rice lines, sorghum lines, Camelina lines, Pongamia lines, Brassica carinata lines, castor lines, or tomato lines.
In yet a further preferred embodiment of the method of 8, the genetically-divergent, heterozygous plant lines are *Jatropha* lines.

In a particularly preferred embodiment of the methods described herein, the *Jatropha* lines are selected from the group consisting of undomesticated *Jatropha*, inbred *Jatropha*, single cross F1 hybrid *Jatropha*, clonally propagated *Jatropha* clones, and a combination of any of the foregoing.

Preferably, in an embodiment of the methods described herein, the target trait is selected from the group consisting of high productivity, synchronous flowering, high oilseed yield, day length insensitivity, growth habit uniformity, abiotic stress tolerance (e.g., cold tolerance, drought tolerance, wet tolerance, salinity tolerance, herbicide tolerance), biotic stress tolerance (e.g., insect tolerance, nematode tolerance, disease resistance including but not limited to resistance to rust caused by *Phakopsora jatrophicola*, root rot caused by *Clitocybe tabescens*, leaf spot caused by *Colletotrichum gloeosporioides*), water use efficiency, nitrogen use efficiency, and combinations of any thereof, or any other agronomically beneficial trait or desirable characteristic as defined herein.

In a preferred embodiment of the methods described herein, the target trait is high productivity.

In another aspect, the present invention provides an essentially homozygous parental plant line, more preferably, an elite plant line produced by of the methods described herein. Thus, seeds, plants, and plant parts of such essentially homozygous parental line are within the scope of the plants and methods described herein.

In a preferred embodiment, the essentially homozygous parental line produced by the methods of the invention is essentially a homozygous *Jatropha* parental line.

In a preferred embodiment, the essentially homozygous parental plant line described is a *Jatropha* parental line.

In another preferred embodiment, the present invention provides a plant of the essentially homozygous parental plant line as described herein.

In another preferred embodiment, the present invention also provides a part of the plant as described herein.

In another preferred embodiment, the part of the plant as described herein is selected from the group consisting of a protoplast, a cell, a tissue, an organ, a cutting, and an explant.

In yet another preferred embodiment, the part of the plant of 16 is selected from the group consisting of an inflorescence, a flower, a sepal, a petal, a pistil, a stigma, a style, an ovary, an ovule, an embryo, a receptacle, a seed, a fruit, a stamen, a filament, an anther, a male
or female gametophyte, a pollen grain, a meristem, a terminal bud, an axillary bud, a leaf, a stem, a root, a tuberous root, a rhizome, a tuber, a stolon, a corm, a bulb, an offset, a cell of said plant in culture, a tissue of said plant in culture, an organ of said plant in culture, and a callus.

In another aspect, the present invention further provides progeny or seed of the plant as described herein.

In another aspect, the present invention further provides a method of producing a hybrid Jatropha seed, comprising crossing a first Jatropha plant of an essentially homozygous Jatropha parental line as described herein with a second Jatropha plant, and producing said hybrid Jatropha seed.

In an embodiment of the method as described herein, the second Jatropha plant is of an essentially homozygous Jatropha parental line.

In a preferred embodiment of the methods described herein, the second essentially homozygous Jatropha parental line is produced by the method of any one of methods described herein.

In another aspect of the present invention, provided is a hybrid Jatropha seed produced by the methods described herein.

In yet another aspect, the present invention further provides a hybrid Jatropha plant grown from seed described herein.

In a preferred embodiment, the hybrid Jatropha plant described herein exhibits an improved target trait.

In a preferred embodiment, the improved target trait of the hybrid Jatropha plant described herein is selected from the group consisting of high productivity, synchronous flowering, high oilseed yield, day length insensitivity, biotic stress tolerance, abiotic stress tolerance, water use efficiency, nitrogen use efficiency, growth habit uniformity, and combinations thereof.

In a preferred embodiment, the improved target trait of the hybrid Jatropha plant described herein is high productivity.

The present invention further provides, in one aspect, a part of the hybrid Jatropha plant as described herein.

In an embodiment, the part of the hybrid Jatropha plant as described herein is selected from the group consisting of a protoplast, a cell, a tissue, an organ, a cutting, and an explant.

In an embodiment, the part of the hybrid Jatropha plant as described herein, is selected from the group consisting of an inflorescence, a flower, a sepal, a petal, a pistil, a stigma, a
style, an ovary, an ovule, an embryo, a receptacle, a seed, a fruit, a stamen, a filament, an anther, a male or female gametophyte, a pollen grain, a meristem, a terminal bud, an axillary bud, a leaf, a stem, a root, a tuberous root, a rhizome, a tuber, a stolon, a corm, a bulb, an offset, a cell of said plant in culture, a tissue of said plant in culture, an organ of said plant in culture, and a callus.

The present invention also provides, in one aspect, progeny or seed of said hybrid *Jatropha* plant described herein.

The present invention further provides oil obtained from the *Jatropha* seed described herein.

The present invention further provides a product produced from the oil of the *Jatropha* seed described herein, wherein said product is selected from the group consisting of biodiesel, biokerosene, hydraulic fluids, dielectric coolants, bioplastics, specialty chemicals, and pharmaceutical intermediates.

In some embodiments, the method for selecting essentially homozygous parental plant lines for the production of a hybrid progeny according to the invention further comprise the steps of enhancing the breeding value of the new essentially homozygous parental lines by intercrossing, selfing and backcrossing while continuously selecting for the genomic regions associated with the agronomically desirable traits. Experimental crosses are done to evaluate general combining abilities (intercrossability).

In addition to the exemplary aspects and embodiments described above, further aspects and embodiments will become apparent by study of the following descriptions. It should be understood, however, that the detailed description and any specific examples provided, while indicating specific embodiments of the invention, are given by way of illustration only, since various changes and modifications within the spirit and scope of the invention will become apparent to those skilled in the art from this detailed description.

**BRIEF DESCRIPTION OF THE FIGURES**

**FIG. 1.** Phenotypic variation in a typical cross between two genetically dissimilar *Jatropha* parents.

**FIG. 2.** Seed yield data from a large number of F1 progeny resulting from crossing of heterozygous parental lines shows a binomial seed yield distribution curve. Both high- and low-yielding transgressive segregants are selected for subsequent genetic marker association studies.
FIG. 3. Identification of fingerprints characteristic of high breeding gain. Fingerprints tightly associated with high performing hybrids, (black bars); fingerprints tightly associated with poorly performing hybrids (gray bars).

FIG. 4. DNA fingerprinting can accelerate selection of inbred and dihaploid parental lines. Genome-wide SNP maps of inbred plants are compared to the consensus SNP map for productive plants and the inbreds most closely approximating the consensus are selected.

DETAILED DESCRIPTION OF THE INVENTION

The present disclosure relates to materials and methods useful for improving the efficacy of a plant breeding program such as, for example, the process of domestication of wild plant populations to accelerate the development of essentially homozygous lines and varieties of plants which in turns are useful for, for example, commercial production of uniform F1 hybrid progeny. In some preferred embodiments, plant breeding programs as described herein are aimed at altering and/or improving phenotypic traits for which associations with genetic markers can be established. The present disclosure further relates to a method for producing hybrid seeds and to hybrid seeds thus produced.

Definitions

Unless otherwise defined, all terms of art, notations and other technical and scientific terms or terminology used herein are intended to have the meanings commonly understood by those of skill in the art to which this invention pertains. Many of the techniques and procedures described or referenced herein are well understood and commonly employed using conventional methodology by those skilled in the art. The following terms are defined for purposes of the invention as described herein. In some cases, terms with commonly understood meanings are defined herein for clarity and/or for ready reference, and the inclusion of such definitions herein should not necessarily be construed to represent a substantial difference over what is generally understood in the art.

The singular forms "a", "an", and "the" include the plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to "a host cell" includes a plurality of such host cells, and a reference to "a stress" is a reference to one or more stresses and equivalents thereof known to those skilled in the art, and so forth.

As used herein, the term "allele" refers to any of one or more alternative forms of a gene locus, all of which alleles relate to one trait or characteristic. In a diploid cell or organism, the two alleles of a given gene occupy corresponding loci on a pair of homologous chromosomes. Each copy may be a distinct allele.
"The term "backcrossing" as used herein refers to the repeated crossing of a hybrid progeny back to the recurrent parents. The parental plant which contributes the gene for the desired characteristic is termed the nonrecurrent or donor parent. This terminology refers to the fact that the nonrecurrent parent is used one time in the backcross protocol and therefore does not recur. The parental plant to which the gene or genes from the nonrecurrent parent are transferred is known as the recurrent parent as it is used for several rounds in the backcrossing protocol. In a typical backcross protocol, the original variety of interest (recurrent parent) is crossed to a second variety (nonrecurrent parent) that carries the single gene of interest to be transferred. The resulting progeny from this cross are then crossed again to the recurrent parent and the process is repeated until a plant is obtained wherein essentially all of the desired morphological and physiological characteristics of the recurrent parent are recovered in the converted plant, in addition to the single gene or a limited number of genes transferred from the nonrecurrent parent. Backcrossing can be used to introduce one or more single locus conversions from one genetic background into another.

As used herein, "plant biomass" refers to the amount of (e.g., measured in grams of air-dry tissue) of a harvestable plant tissue produced from the plant in a growing season, which could also determine or affect the plant yield or the yield per growing area. Non-limiting examples of such harvestable plant tissues include leaves, stems, and reproductive structures, or all plant tissues such as leaves, stems, roots, and reproductive structures.

The term "crossing" as used herein refers to the fertilization of female plants (or gametes) by male plants (or gametes). The term "gamete" refers to the haploid reproductive cell (egg or sperm) produced in plants by mitosis from a gametophyte and involved in sexual reproduction, during which two gametes of opposite sex fuse to form a diploid zygote. The term generally includes reference to a pollen (including the sperm cell) and an ovule (including the ovum). "Crossing" therefore generally refers to the fertilization of ovules of one individual with pollen from another individual, whereas "selfing" refers to the fertilization of ovules of an individual with pollen from the same individual. Crossing is widely used in plant breeding and results in a mix of genomic information between the two plants crossed one chromosome from the mother and one chromosome from the father. This will result in a new combination of genetically inherited traits. Usually, the progeny of a crossing is designated as: "Fl". If the Fl is not uniform (segregates) it is usually designated as "Fl population". "Selfing" of a homozygous plant will usually result in a genetic identical plant since there is no genetic variation. "Selfing" of an Fl will result in an offspring that segregates for all traits that have heterozygotic loci in the Fl. Such offspring is designated: "F2" or "F2 population".
The term "polycross" or "polycrossing" refers to a cross used in selective plant breeding which is a method of mass experimental crossbreeding. It involves finding clones of strains which, upon crossbreeding with other clones or strains of the same species, yield the most productive plants. The resulting plants are used in developing a new "synthetic variety." This method is commonly used in the selective breeding of plants that can be successfully cloned (perennial herbs and annuals and biennials that propagate vegetatively). The term "intercrossable", as used herein, refers to the ability to yield progeny plants after making crosses between parent plants.

As used herein, the term "cross-pollination" refers to fertilization by the union of two gametes from different plants. A plant is cross-pollinated if the pollen comes from a flower on a different plant from a different family or line. Cross-pollination does not include sib- and self-pollination.

As used herein, a "cultivar" or a "variety" refers to a group of similar plants that belong to the same species and that, by structural features and performance, may be distinguished from other varieties within the same species. Two essential characteristics of a variety are identity and reproducibility. Identity is necessary so that the variety may be recognized and distinguished from other varieties within the crop species. The distinguishing features may be morphological characteristics, molecular markers, color markings, physiological functions, disease reaction, or performance. Most agricultural varieties are pure for the characteristic or for those characteristics that identify the variety; per se. Reproducibility is needed in order that the characteristic(s) by which the variety is identified will be reproduced in the progeny. For the purpose of this disclosure, therefore, the terms "cultivar" and "variety" are used interchangeably to refer to a group of plants within a species (here, *Jatropha curcas*) that share certain constant characters which separate them from the typical form and from other possible varieties within that species. While possessing at least the distinctive trait, a "variety" of the invention also may be characterized by a substantial amount of overall variation between individuals within the variety, based primarily on the Mendelian segregation of traits among the progeny of succeeding generations. On the other hand, "cultivar" or "variety" also can denote a clone, since a *Jatropha curcas* cultivar may individually be reproduced asexually, via stem cuttings, and all of the clones would be essentially identical genetically.

A "line", as used herein, refers to a population of plants derived from a single cross, backcross or selfing. The individual offspring plants are not necessarily identical to one another. As distinguished from a "variety," a "line" is displays less variation between individuals, generally (although not exclusively) by virtue of several generations of self-
pollination. For purposes of this disclosure, a "line" is defined sufficiently broadly to include a group of plants vegetatively propagated from a single parent plant, using stem cuttings or tissue culture techniques.

The term "breeding line" as used herein, refers to a line of a cultivated crop having commercially valuable or agronomically desirable characteristics, as opposed to wild varieties or landraces. The term includes reference to an elite breeding line or elite line, which represents an essentially homozygous, usually inbred, line of plants used to produce commercial F1 hybrids. An elite breeding line is obtained by breeding and selection for superior agronomic performance comprising a multitude of agronomically desirable traits. An elite plant is any plant from an elite line. Elite breeding lines are essentially homozygous and are preferably inbred lines.

As used herein, altered flowering time may mean either that: (a) flowering time of a plant is earlier than a wild-type plant (i.e., early flowering) or (b) the flowering time is later than a wild-type plant (i.e., late flowering). Early Flowering: Plant species vary in the temporal lengths of their life cycles. Plants can have life cycles that may be completed within one year or span across several years. Plants generally flower late in their life cycle, after embryogenesis, seedling development and a period of vegetative growth (Walbot (1985) Trends Genet. 1:165-169). Flowering time in plants is influenced by many endogenous and environmental factors, including gibberellin biosynthesis and signaling, autonomous controls, light quality and intensity, photoperiod, temperature and availability of nutrients (Garner and Allard (1920) J. Agric. Res. 18:553-606; Bernier (1988) Annu. Rev. Plant Physiol. 39:175-219; Millar (1999) New Phytol. 14:175-197; Battey (2000) J. Exp. Bot. 51:1769-1780; Samach and Coupland (2000) Bioessays, 22:38-47). Early flowering may mean: (a) that the plant has begun to flower at a time statistically significantly earlier than another plant or plants of a different variety grown under the same conditions because the transition from the vegetative pre-flowering phase to the reproduction phase occurs earlier; or (b) that the plant has begun to flower at a time statistically significantly earlier than another plant or plants of a different variety grown under the same conditions because the growth rate of the plant prior to flowering and/or in the entire life cycle has been enhanced. Early flowering may also be described as a plant flowering at a moment in its life cycle that is at least 1% to 10% earlier in the plant's life cycle compared to another plant or plants of a different variety grown under identical conditions. Alternatively, the plant may begin to flower at a moment in the plant's life cycle that is at least 10% to 25% earlier or at least 25% to 50% earlier or at least 50% to 99% earlier. "Late flowering": In contrast to early flowering, late flowering may mean: (a) that the plant has begun to flower at
a time statistically significantly later than another plant or plants of a different variety, grown under the same conditions because the transition from the vegetative, pre-flowering phase to the reproduction phase occurs later; or that the plant has begun to flower at a time statistically significantly later than another plant or plants of a different variety, grown under the same conditions because the growth rate of the plant prior to flowering and/or in the entire life cycle has been decreased. Late flowering can also be described as a plant flowering at a moment in its life cycle that is at least 1% and 10% later in the plant's life cycle compared to a corresponding wild-type plant grown under identical conditions. Alternatively, the plant may begin to flower at a moment in the plant's life cycle that is at least 10% to 25% later or at least 25% to 50% later or at least 50% to 99% later.

As used herein, the expressions "molecular genetic marker" or short "genetic marker" are used interchangeably herein and refer to a region of a nucleotide sequence (e.g., in a chromosome) that is subject to variability (i.e., the region can be polymorphic for a variety of alleles). Genetic markers are typically used in methods for visualizing differences in characteristics of nucleic acid sequences. Examples of such indicators are restriction fragment length polymorphism (RFLP) markers, amplified fragment length polymorphism (AFLP) markers, single nucleotide polymorphisms (SNPs), insertion/deletion (INDEL) mutations, simple sequence repeats (SSRs or microsatellite) markers, sequence-characterized amplified regions (SCARs), cleaved amplified polymorphic sequence (CAPS) markers, restriction fragment length polymorphisms (RFLPs), randomly amplified polymorphic DNAs (RAPDs), isozyme markers, arbitrarily primed polymerase chain reaction (AP-PCR), DNA amplification fingerprinting (DAF) or combinations of the markers described herein, which define a specific genetic and chromosomal region or chromosomal location.

Genetic markers can be used during the breeding process for the selection of qualitative and/or quantitative traits. For example, markers closely linked to alleles or markers containing sequences within the actual alleles of interest can be used to select plants that contain the alleles of interest during a backcrossing breeding program. The markers can also be used to select for the genome of the recurrent parent and against the genome of the donor parent. Using this procedure can minimize the amount of genome from the donor parent that remains in the selected plants. It can also be used to reduce the number of crosses back to the recurrent parent needed in a backcrossing program. The use of genetic markers in the selection process is often called genetic marker enhanced selection or marker-assisted selection (MAS). Genetic molecular markers may also be used to identify and exclude certain sources of germplasm as
parental varieties or ancestors of a plant by providing a means of tracking genetic profiles through crosses.

As used herein, the term "genotype" refers to the genetic constitution of a cell or organism.

As used herein, the terms "haploid" and "doubled-haploid" refers to a cell or organism having one set of the two sets of chromosomes in a diploid. Doubled haploids are plants that have two copies of each chromosome, (2n), like diploids. However, they differ from diploids in that they were created from a single grain of pollen, an ovum, or indeterminate gametes that are cultured. Their chromosomes doubled through chemical means, and the cultured tissue grown into a plant. The haploid genome of the gametes, when doubled, produced a plant with a complete genome, with two identical copies of every gene. Thus, double haploids are homozygous at every locus, and can have highly variable phenotypes. Double haploids have been made for many plant species to assist in breeding experiments.

For the purpose of this invention, the term "homologous" is used to denote a characteristic of a polynucleotide sequence exhibiting at least about 70% sequence identity compared to a reference polynucleotide sequence, more preferably at least about 80-85% sequence identity, more preferably at least about 85%-90% sequence identity, more preferably at least about 90%-95% sequence identity, more preferably at least about 96%, 97%, 98%, or 99% sequence identity, and most preferably about 100% sequence identity as compared to a reference sequence. The reference sequence may be a subset of a large sequence, such as a portion of a gene or a region of a chromosome or genome. Sequence homology can be determined using any techniques or homology comparison software, including for example, the BLASTN software of the National Center for Biotechnology Information (NCBI). DNA sequences that are homologous can also be determined in a Southern hybridization experiment under, for example, stringent conditions, as defined for that particular system. Defining appropriate hybridization conditions is within the skill of the art. See, e.g., Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Second Edition, (1989) Cold Spring Harbor, N.Y.; *DNA Cloning*, Vols. I & II, D.N. Glover, ed. (1985); *Nucleic Acid Hybridization: A Practical Approach*, editors B.D. Hames and S.J. Higgins, (1985) Oxford; Washington, DC; IRL Press.

It is understood that homologous sequence may accommodate insertions, deletions, and substitutions in the nucleotide sequence. Thus, linear sequence of nucleotides can be essentially identical even if some of the nucleotides residues do not precisely correspond or align.
As used herein, the term "homozygous" means a genetic condition existing when identical alleles reside at corresponding loci on homologous chromosomes. Homozygosity levels are average values for the population, and refer preferably to those loci at which the parental genomes are identical.

The expression "essentially homozygous line" refers to a plant line having a level of homozygosity of at least 90%, preferably at least 95%, preferably at least 96%, more preferably at least 97%, more preferably still at least 98%, and most preferably at least 99% or at least 100% homozygosity when testing at least 50, preferably at least 100, preferably at least 300, more preferably at least 500, more preferably at least 1,000; more preferably at least 2,000; and most preferably at least 10,000 loci. In some preferred embodiments, the homozygosity level is determined using molecular methods.

As used herein, the term "heterozygous" means a genetic condition existing when different alleles reside at corresponding loci on homologous chromosomes. The expression "heterozygous line" merely reflects that the line is not an "essentially homozygous line" as used herein.

Hybrid. As used herein, the term "hybrid" means any offspring of a cross between two genetically non-identical individuals. The parental plants may be related, as in production of a modified single cross, or unrelated. F1 hybrid, as used herein, refers to the first generation progeny of the cross of two genetically dissimilar plants.

As used herein, the terms "introgressing", "introgress" and "introgressed" refer to both a natural and artificial process whereby individual genes or entire chromosomes are moved from one individual, species, variety or cultivar into the genome of another individual, species, variety or cultivar, by crossing those individuals, species, varieties or cultivars. In plant breeding, the process usually involves selfing or backcrossing to the recurrent parent to provide for an increasingly homozygous plant having essentially the characteristics of the recurrent parent in addition to the introgressed gene or trait.

As used herein, the term "linkage" refers to a phenomenon wherein alleles on the same chromosome tend to segregate together more often than expected by chance if their transmission was independent.

As used herein, the term "linkage disequilibrium" refers to a phenomenon wherein alleles tend to remain together in linkage groups when segregating from parents to offspring, with a greater frequency than expected from their individual frequencies.

A "locus" is defined herein as the position that a given gene occupies on a chromosome of a given plant species. A locus confers one or more traits such as, for example, male sterility,
female-only flower, herbicide tolerance, pest resistance, disease resistance, synchronous germination, synchronous flowering, early flowering, improved plant yield and/or fruit yield, modified plant architecture, abiotic stress tolerance, modified fatty acid metabolism, modified oil content, modified carbohydrate metabolism, and modified protein metabolism. The trait may be, for example, conferred by a naturally occurring gene introduced into the genome of the variety by backcrossing, a natural or induced mutation, or a transgene introduced through genetic transformation techniques. A locus may comprise one or more alleles integrated at a single chromosomal location.

As used herein, the phrase "oil content" as used herein refers to the amount of lipids in a given plant organ, such as the seeds (seed oil content) and is typically expressed as percentage of dry weight (for example at 10 % humidity of seeds) or wet weight.

It should be noted that oil content is affected by intrinsic oil production of a tissue (e.g., seed, fruit), as well as the mass or size of the oil-producing tissue per plant or per growth period. In one embodiment, increase in oil content of the plant can be achieved by increasing the size/mass of a plant's tissue(s) which comprise oil per growth period. Thus, increased oil content of a plant can be achieved by increasing the yield, growth rate, biomass and vigor of the plant.

As used herein, the term "phenotype": The detectable characteristics of a cell or organism, which characteristics are the manifestation of gene expression.

The term "plant" refers to any living organism belonging to the kingdom Plantae. As used herein, the term "plant" includes reference to an immature or mature whole plant, including a plant from which seed or anther have been removed. A seed or embryo that will produce the plant is also considered to be the plant.

A plant characteristic can be a morphological, physiological, agronomic, or genetic feature of a plant.

As used herein, "plant growth" This term refers to the process by which plants increase in size and mass. The increase in the number and size of plant organs is directly associated with an increase in cell numbers and/or cell size, which involves cell division, growth, expansion and differentiation. Plants utilize sunlight, water, carbon dioxide and minerals in biosynthesis to provide energy and substances required for growth. Plant growth can be generally divided into vegetative and reproductive growth in the life cycle.

Plant height is taken from the top of the soil to the top node of the plant and is measured in centimeters.
The term "plant part" refers to any part of a plant including, but not limited to, organelles, single cells and cell tissues such as plant cells that are intact in plants, cell clumps and tissue cultures from which *Jatropha* plants can be regenerated. Examples of plant parts include, but are not limited to, single cells and tissues from pollen, ovules, leaves, embryos, roots, root tips, tubers, anthers, flowers, fruits, stems shoots, and seeds; as well as pollen, ovules, leaves, embryos, roots, root tips, anthers, flowers, fruits, stems, shoots, scions, rootstocks, seeds, tubers, protoplasts, calli, and the like. The two main parts of plants grown in some sort of media, such as soil, are often referred to as the "above-ground" part, also often referred to as the "shoots", and the "below-ground" part, also often referred to as the "roots".

As used herein, the term "progeny" means (a) genetic descendant(s) or offspring. Progeny includes descendants of a particular plant or plant line. Progeny of a plant according to the present invention include seeds formed on F1, F2, F3, F4, F5, F6 and subsequent generation plants, or seeds formed on BC1, BC2, BC3, and subsequent generation plants, or seeds formed on F1BC1, F1BC2, F1BC3, and subsequent generation plants. The designation F1 refers to the progeny of a cross between two parents that are genetically distinct. The designations F2, F3, F4, F5 and F6 refer to subsequent generations of self- or sib-pollinated progeny of an F1 plant.

Quantitative Trait Loci (QTL): Quantitative trait loci (QTL) refer to genetic loci that control to some degree numerically representable traits that are usually continuously distributed.

As used herein, the terms "resistance" and "tolerance" are used interchangeably to describe a plant having the ability to prevent, decrease, or repair the injury induced by a specified biotic or abiotic stress on a plant or a plant population such as insect pest, pathogenic disease, abiotic influence, or environmental condition. These terms are also used to describe plants showing some stress symptoms but that are still able to produce marketable product with an acceptable yield. Some plants that are referred to as resistant or tolerant are only so in the sense that they may still produce a crop, even though the plants are stunted and the yield is reduced.

As used herein, the term "regeneration" refers to the development of a plant from tissue culture.

"Selfing" refers to the manifestation of the process of "self-pollination", which in turn refers to the transfer of pollen from the anther of a flower to the stigma of the same flower or different flowers on the same plant. The term "selfing" therefore refers to the process of self-fertilization wherein an individual is pollinated or fertilized with its own pollen. Repeated selfing eventually results in homozygous offspring.
As used herein, the term, "Single Locus Converted (Conversion) Plant" refers to plants which are developed by a plant breeding technique called backcrossing, wherein essentially all of the morphological and physiological characteristics of *ajatropa* variety are recovered in addition to the characteristics of the single locus transferred into the variety via the backcrossing technique and/or by genetic transformation.

As used herein, the term "tissue culture" refers to a composition comprising isolated plant cells of the same or a different type or a collection of such cells organized into parts of a plant, in which the cells are propagated in a nutrient medium under controlled conditions.

As used herein, the term "transgressive segregant", as used herein in reference to studies of quantitative traits in segregating hybrid populations, refers to offspring (e.g., created within a breeding program) that possess significantly different traits/phenotypes than their parents, *i.e.* offspring having phenotypes that are extreme relative to those of either parental lines. As such, in segregating hybrid generations, transgressive segregants are the fraction of individuals that exceed parental phenotypic values in either a negative or positive direction. Transgressive segregation, in genetics, is the formation of extreme phenotypes, or transgressive phenotypes, observed in segregated hybrid populations compared to phenotypes observed in the parental lines. Typically, the frequency of transgression within a segregating hybrid population is positively correlated with the genetic divergence of the parental lines used to create the segregating population.

As used herein, "vegetative propagation" refers to asexual propagation of the plant that is accomplished by taking and propagating cuttings, by grafting or budding, by layering, by division of plants, or by separation of specialized structure, such as stem, roots, tubers, rhizomes, or bulbs.

As used herein the phrase "plant vigor" refers to the amount (measured by weight) of tissue produced by the plant in a given time. Hence increased vigor could determine or affect the plant yield or the yield per growing time or growing area. In addition, early vigor (seed and/or seedling) often results in improved field stand establishment. As used herein, stand establishment refers to the survivability and density of areas of land newly planted with jatropha, typically by seed or stem propagation.

As used herein the phrase "plant yield" refers to the amount (as determined by, *e.g.* weight or size) or quantity (numbers) of tissues or organs produced per plant or per growing season. Hence increased yield could affect the economic benefit one can obtain from the plant in a certain growing area and/or growing time. It should be noted that a plant yield can be affected by various parameters including, but not limited to, plant biomass; plant vigor; growth
rate; seed yield; seed quantity; seed quality; oil yield; oil content; starch and/or protein in harvested organs (e.g., seeds or vegetative parts of the plant); number of flowers (florets) per inflorescence; harvest index; number of plants grown per area; number and size of harvested organs per plant and per area; number of plants per growing area (density); number of harvested organs in field; total leaf area; carbon assimilation and carbon partitioning (the distribution/ allocation of carbon within the plant); resistance to shade; number of harvestable organs (e.g., seeds), weight per seed; and modified plant architecture.

As used herein the phrase "seed yield" refers to the number or weight of the seeds per plant, per growing season, or per growing area or to the weight of a single seed. Hence seed yield can be affected by seed dimensions (e.g., length, width, perimeter, area and/or volume), number of (filled) seeds and seed filling rate and by seed oil content. Hence an increase of seed yield per plant could affect the economic benefit one can obtain from the plant in a certain growing area and/or growing time; and an increase of seed yield per growing area could be achieved by increasing seed yield per plant, and/or by increasing number of plants grown on the same given area.

*Jatropha curcas*

*Jatropha curcas* is a flowering plant species in the spurge family, *Euphorbiaceae*, and is thus closely related to other important cultivated plants such as rubber and castor. The genus *Jatropha* comprises over 170 species. The commonly occurring species are *J. curcas*, *J. glandulifera*, *J. gossypifolia*, *J. multifida*, *J. nana*, *J. panduraefolia*, *J. villosa* and *J. podagrica*. While many of the *Jatropha* species are used for ornamental purposes, the two species *J. curcas* and *J. glandulifera* are commonly known as oil yielding species. *Jatropha curcas* is widely thought to have originated from tropical America, but is now found abundantly in many parts of the tropics and sub-tropics in Africa and Asia. Common names include barbados nut, physic nut, purging nut, Curcas bean, and noix-de-medicine.

The *Jatropha* plant is a small tree or large shrub that can reach a height of up to 5 meter. *Jatropha curcas* reportedly can grow on a mere 250 mm (10 inches) of rain a year. Ploughing and planting are typically not needed regularly, as this shrub has a life expectancy of approximately forty years. The species is monoecious, with inflorescences formed terminally on branches, possessing main and co-inflorescences with paracladia.

*Jatropha curcas* can be propagated vegetatively, as well as by seeds. Vegetatively, this crop can be propagated by stem cuttings, grafting, and budding as well as by air layering techniques. Sexually, it is a cross-pollinating species. The problems of low viability and recalcitrant nature of oil seeds may limit the sexual propagation. Complete germination is
typically achieved within 9 days. The flowers typically develop terminally, \textit{i.e.}, at the end of a stem, so a good ramification (plants presenting many branches) produces a greater amount of fruits. Another productivity factor is the ratio between female and male flowers within an inflorescence - more female flowers mean more fruits. Although the number of male and female flowers per inflorescence varies greatly among different observations, the male-to-female flower ratio is relatively similar.

As with many members of the family \textit{Euphorbiaceae}, \textit{Jatropha} contains several toxic compounds. The toxicity of the seeds and leaves is due to various components, including a toxic protein (curcin), hydrocyanic acid, and diterpene esters. Curcin is related to ricin, the toxic protein of the castor bean (\textit{Ricinus communis}), but curcin lacks the lectin domain that makes ricin so extremely toxic. Curcin is reported to hinder protein synthesis \textit{in vitro}. Diterpenes have been isolated from seeds and roots and reportedly promote skin tumors when combined with normally sub carcinogenic levels of known carcinogens.

In Asia and Africa, the \textit{Jatropha} plant is mainly grown as a living fence around homesteads, gardens and fields, since it is poisonous and not browsed by animals. Each and every part of the tree from roots to the leaves can be used for various purposes, \textit{e.g.}, to make antibiotics, medicine for skin diseases treatment and others. Different parts of \textit{Jatropha} plants have various medicinal uses especially in nutraceuticals, pharmaceutical, dermatological, and personal care products. The latex from \textit{Jatropha curcas} is reported to have an anticancer properties associated with various alkaloids including Jatrophone, \textit{Jatropham} and curcain. The seeds of \textit{Jatropha} contain viscous oil that is reported to possess insecticidal, mollucicidal, fungicidal and nematicidal properties. In many developing countries, the residue of oil extraction (seed cake) can also be processed and used as biomass feedstock to power electricity plants or used as organic fertilizer, because it contains nitrogen, phosphorous and potassium. Seed cake can also be used as potential feedstock in gasification plants. Its reported value for Gross Calorific Value (GCV) is 18 MJ/kg, and its humidity content is approximately 8%, making \textit{Jatropha} seed cake a possible biomass feedstock for syngas production. \textit{Jatropha} can also be used as animal feed, biopesticide and rodent repellent. \textit{Jatropha} oil is a strong purgative; therefore in countries such as India, it is widely used as an antiseptic for cough, skin diseases and as a pain reliever in rheumatism. Refining crude \textit{Jatropha} oil into biofuel products produces glycerine as by-product, which is in great demand as a raw material for cosmetic, medicine and food product industries. \textit{Jatropha} seed oil can also be used for manufacture of candles and soap, in cosmetics industry, in formulating lubricants, softeners, dyeing assistants, making Turkey red oil and the adulteration of olive oil, and as a diesel/paraffin substitute or
extender. This latter use has important implications for meeting the demand for rural energy services and also exploring practical substitutes for fossil fuels to counter greenhouse gas accumulation in the atmosphere. These characteristics along with its versatility make it of vital importance to developing countries.

*Jatropha* seed production is typically around 0.5 to 1.5 tons per hectare, ranging from as littles as zero production per hectare in first year to over 1.5 tons per hectare after 3 years. If planted in hedges, the reported productivity of *Jatropha* is from 0.8 kg to 1.0 kg of seed per meter of live fence. The seeds contain semi-dry oil which has been found useful for several medicinal and veterinary purposes. Typically, the oil content is 25-40% in the seeds and 50-60% in the kernel, which is extracted mainly from the endosperm tissues. The oil contains about 20% saturated fatty acids, and about 80% unsaturated fatty acids. *Jatropha* seed oil belongs to the oleic or linoleic acid group, to which the majority of vegetable oils belong. Linoleic acid (C18:2) and oleic acid (C18:1) together account for up to 80% of the oil composition. Palmitic acid (C16:0) and stearic acid (C18:0) are other fatty acids present in this oil.

The demand for oil and fats is expected to increase dramatically with the increase in world population. In the last few years, the non-edible oils market has developed largely due to the increasing demand for biofuels. Among the non-edible oils, *Jatropha* oil has been acknowledged for its physical and physicochemical properties, found suitable for use as biodiesel feedstock and industrial application. For example, crushing *Jatropha* seeds produces *Jatropha* oil, which can be processed to produce a biodiesel that meets the restrictive specifications of EU biodiesel fuel standard (EN 14214), which reportedly can be used in a standard diesel car. Moreover *Jatropha* oil is reported to fit well within the specification requirements of the main providers of endothermic engines, which generate electricity and heat to run straight on vegetable oil. Recently, *Jatropha* oil has been tested in Air New Zealand’s Boeing 747, running one of its four Rolls-Royce engines on a 50:50 blend of *Jatropha* oil and jet A-1 fuel. Subsequently in January 2009, Air New Zealand and Continental Airlines have run test flights further demonstrating the viability of *Jatropha* oil as a jet fuel.

In addition, *Jatropha curcas* is gaining commercial importance as a biodiesel plant, and has attracted attention in terms of developing the plant in wastelands and dry lands. The increasing interest in oil from *Jatropha curcas* has generated enormous pressure to supply enough seeds that are homogenous and productive enough for plantation. As the need for biofuels increases, the higher productivity of *Jatropha* makes it even more attractive as a target for biotechnological improvement. There is a need to increase the quality and yield of *Jatropha*
oil and to rapidly develop useful characteristics when required, such as traits that enable the plant to grow in more arid locations, with higher oil yield and lower height increments. *Jatropha* is increasingly considered an important and valuable crop. Thus a continuing goal of plant breeders is to develop stable high yielding jatropha cultivars that are agronomically sound. The reasons for this goal are obvious to maximize the amount of seed and plant parts produced per unit land area.

However, although *Jatropha* is known for having wide adaptability and plethora of uses, the full potential of *Jatropha* is far from being realized because of several reasons. Relatively little research has been done to improve *Jatropha* varieties because the vast majority of *Jatropha* is produced by subsistent farmers. The crop's growth is therefore limited by insect pests, disease and weeds, rather than the plant's inherent ability. Several pests and diseases have been reported for *Jatropha*, e.g. the seed-feeding Scutelleridae, *Agonosoma trilineatum*; the scutellarid bug, *Scutellera nobilis*; and the inflorescence and capsule-borer, *Pempelia morosalis*. Typically, fruits of *Jatropha* in one brunch will ripe at different times, which cause difficulties during harvesting and leads to higher labor inputs. Therefore, new *Jatropha* varieties with uniform fruit maturation and ripening would be desirable and more suitable for large-scale mechanical harvesting methods. In recent years, research in *Jatropha* germplasm has increased, and a host of different *Jatropha* accessions from many different geographic areas are being assessed for desirable traits. However, very few of the desirable traits have been introgressed into commercial planting materials. In addition, given the previously limited knowledge about *Jatropha* genetics, a number of steps are required in order to improve germplasm domestication programs.

Therefore, there is a great potential for improvements using conventional breeding techniques as well as marker-assisted breeding methods. In addition, modification of *Jatropha* products *in situ*, or with modern genomics tools as described herein, adds significant value to the plantations and provides a wide range of differentiated products for a wide variety of applications. Success of such efforts is greatly enhanced by the availability of the *Jatropha* genome sequence, and the identification of relevant genes of interest. Improvements in growth and oil production, production of non-toxic oils and plant products, recovery of higher and more economical yields of renewable fuels, together with environmental benefits will increase the returns from *Jatropha* plantations on eroded land, and encourage adoption by small and marginal farmers in the tropics.

It is therefore an objective of the present invention to provide means and methods for the production of elite breeding plant lines and for the production of essentially uniform hybrid...
seed from which plants can be grown that exhibit agronomically relevant traits such as, for example, enhanced growth and oil production properties.

**Detailed Description of the Inventive Methods**

The production of uniform hybrid varieties generally requires the development of homozygous inbred plants, followed by the crossing of these inbred plants, and the evaluation of the crosses. Plants that have been self-pollinated and selected for type over many generations become homozygous at almost all gene loci and produce a uniform population of true breeding progeny, which is a homozygous plant. A cross between two such homozygous, near-isogenic plants of different varieties typically produces a uniform population of hybrid plants that display the same allelic heterozygosity at many loci.

In the case of *Jatropha curcas*, perhaps the biggest challenge in developing commercial grade *jatropha* hybrids is the undomesticated nature of the species. Most of the jatropha parental lines currently used in large-scale production of jatropha hybrids are heterozygous at many loci and thus lack genetic uniformity. As such, a sexual cross between two genetically dissimilar parents typically results in a heterogeneous collection of F1 hybrids, with each individual plant from the same cross exhibiting a unique allelic segregation event. Applicants have observed not only genotypic segregation among F1 individuals, but also dramatic phenotypic differences. Further, the process of inbreeding of parent lines is often time consuming, especially since the phenology of jatropha parental lines does not necessarily correlate well with yields of the F1 hybrid progeny, a similar conundrum is well documented in maize and other domesticated crops where hybridization is involved.

The detection and exploitation of genetic variation have always been an integral part of plant breeding. In particular, DNA-based molecular markers are useful for detecting the genetic variation available in germplasm collections and/or breeding lines. These markers have been used extensively for the development of saturated molecular genetic maps and physical maps and for the identification of genes or quantitative trait loci (QTLs) controlling traits of economic importance which are in turns can be used for marker-assisted selection. During the past two decades, a number of new next-generation sequencing (NGS) technologies have been developed and subsequently deployed to generate DNA sequence data inexpensively and at a rate that is several orders of magnitude faster than that of traditional technologies. As a result, genomics-assisted breeding approaches have greatly advanced with the increasing availability of genome and transcriptome sequence data for several model plant and crop species. Examples of such genomics-assisted breeding approaches include marker assisted selection (MAS), genome selection (GS), and genome wide association studies (GWAS) (for review see,
e.g., Varshney et al., *Next-generation sequencing technologies and their implications for crop genetics and breeding.* Trends Biotechnol. 27(9):522-530, 2009. GWAS methods in particular have attracted significant interest in plant breeders. The GWAS method typically involves an examination of many common genetic variants in different individuals to see if any variant is associated with a trait. GWAS technology focuses on associations between single-nucleotide polymorphisms (SNPs) and specific traits such as increased yield or disease resistance. These studies typically compare the DNA sequence information of two groups of plants that either displays or lacks the target trait (*i.e.*, controls). Each plant gives a sample of DNA, from which millions of genetic variants are read using SNP arrays. If one type of the variant (one allele) is more frequent in plants with the disease resistance, the SNP is said to be "associated" with the disease resistance. The associated SNPs are then considered to mark a region of the genome which influences the risk of disease, for example. In contrast to methods which specifically test one or a few genetic regions, the GWAS studies investigate the entire genome. This approach is therefore considered to be non-candidate-driven in contrast to gene-specific candidate-driven studies. As such, GWAS identifies SNPs and other variants in DNA which are associated with a trait, but cannot on their own specify which genes or genetic elements are causal.

The breeding methods disclosed herein differ fundamentally from the MAS, GS and GWAS methods described above. These described methods are typically deployed in breeding programs involving isogenic plants whereas the method of the present invention is intended for use in breeding programs with wild plants and landraces which exhibit high allelic heterozygosity. For example, in GS one evaluates both phenotypically and genotypically a training population of ~500 individuals to compute the genome breeding gain, which is predictive and can be used to accelerate breeding rates. MAS methods typically relate to one or several SNPs (DNA markers) associated with a particular trait. However since most agronomically relevant traits are a result of complex QTLs, MAS has not significantly delivered value to breeders. While GWAS does look at genome wide markers and attempts to associate them with traits, this method is not used to advance the development of parental lines for improved hybrid productivity, as described herein. In contrast, the breeding method in accordance to the present invention utilizes high-throughput genomics technologies to identify genome-wide SNPs (SNP fingerprints) which are associated with agronomically important traits, *e.g.* high productivity, in individual F1 plants resulting from crosses between heterozygous parental lines. Methods for determining the statistical significance of an association between a phenotype and a fingerprint or a set of fingerprints (of a given genotype)
may be determined by any statistical test known in the art and with any accepted threshold of statistical significance being required. The application of particular methods and thresholds of significance are also well within the skill of the ordinary practitioner of the art.

The SNP fingerprints are then mapped back to the parental lines and used to identify inbred or dihaploid plants which harbor the collection of SNP fingerprints associated with the productive transgressive segregant individuals. In this way, breeders can use the method disclosed herein to identify the inbred plants with allelic segregation events predicted, when crossed will produce phenotypes matching the originally observed in the transgressive segregants.

Thus, in one aspect of the present invention, there is provided a novel breeding method in which molecular genetic tools are developed to accelerate the domestication and improvement of wild populations by allowing the reconstruction of relatedness information from mixtures of genetically distinct families. This novel method circumvents the main technical obstacles to rational selective inbreeding by allowing breeders to avoid phenology as the only basis for selection of inbred plants to take forward through the domestication process. While it has been shown that genetically divergent parental lines, when crossed can result in F1 progeny exhibiting hybrid vigor (heterosis), it is also known in the art, that phenotypic characteristics of the inbred parental lines do not necessarily correlate with the productivity of the F1 progeny. The underlying mechanisms of heterosis are complex and the genetic basis remains a scientific mystery. The GWAD method generates genome fingerprints associated with productive F1 individuals, which are then mapped back to the parental lines, thus enabling the use of DNA fingerprinting to assist in the selection of inbred parental plants to be taken forward in the domestication process.,

In one aspect, Applicants contemplate utilizing genome-wide survey of SNPs (SNP fingerprints) to assist in identifying specific alleles of each parental genome that should be captured during parental line inbreeding as well as during subsequent selection of the inbred offspring possessing the alleles identified in such manner. Subsequently, consensus 'high yield' and 'low yield' SNP fingerprints can be identified by applying statistical analysis and algorithms to the SNP fingerprints of the top and bottom performing individuals among hybrid progeny derived from a single bi-parental cross. The designation of 'high yield' marker refers to a targeted SNP marker or markers that is found associated with high performing plants, and not found in association with the low yielding plants from the same cross. Similarly, a consensus 'low yielding' SNP fingerprint refers to a SNP marker that is found associated only with low yielding plants and not with high yielding plants from the same cross. The inheritance
pattern of the 'high-yielding' SNP fingerprints in the offspring can be tracked back to the male and female parental lines by comparing the parental SNP fingerprints to those of their high-yielding offspring. Such comparisons allows for identifying targeted SNPs which, if found in any one of the parental lines, have an increased likelihood of resulting in high-performing offspring. These results can then be used to select inbred offspring of the foregoing parental lines to be included the next round(s) of genetic improvement in the domestication (inbreeding) process. Since, the outcome of the parental line crosses can be assessed during the years following the selections, the algorithms developed for GWAD can be iteratively improved and refined, based on actual phenotypic data.

As it will be immediately appreciated by one skilled in the art, the methodology disclosed herein is particularly useful in accelerating the domestication process of wild populations. For example, a typical inbreeding cycle for *Jatropha curcas* requires about 1 year, with a minimum of 5 cycles necessary to achieve a good level of uniformity; however, it is very difficult to assess parental line productivity based on selection of the inbreds. Applying the SNP fingerprint strategy disclosed herein for plant domestication could significantly improve the capacity of the parental lines to produce high yielding hybrid progeny exhibiting high uniformity, because at each cycle of inbreeding, individual plant chosen to take forward will retain SNP markers and genome-wide fingerprints associated with high yielding hybrids. Moreover, the methodologies developed through this study on *jatropha* may have important ramifications for other plant and animal domestication programs.

In another aspect, the methodologies disclosed herein can be used for rapidly improving heterozygous parental lines towards maximizing productivity in hybrid progeny. Whole genome sequencing and genotyping by sequencing (GBS) can be used to identify non-homozygous portions of the parental genomes that have come together in the hybrid progeny to provide high productivity. Sequence information generated in such manner can then be used during the inbreeding process to capture those important regions/portsions from each of the two parents. As a result, the domestication of the foregoing parental lines can be rapidly advanced towards high productivity and uniformity in the hybrid progeny.

The breeding methods disclosed herein have a number of applications, *e.g.*, in applied breeding programs. For example, the methods can be used to predict the phenotypic performance of plant hybrid progeny, *e.g.*, a single cross hybrid produced (actually or hypothetically) by crossing a given pair of inbred lines of known marker genotypes. Similarly, by allowing prediction of phenotypic performance of the potential progeny from a cross, the methods can facilitate selection of plants (*e.g.*, inbred plants, hybrid plants, *etc.*) for use as
parents in one or more crosses; the methods permit selection of parental plants whose offspring have the highest probability of possessing the desired phenotype. This new breeding method can also be deployed in the domestication of other wild populations by associating genome-wide markers with individuals expressing desirable domestication traits and possibly to create new domesticated and cultivated forms of plants. Moreover, the breeding methodologies described herein may also have utility in crop re-domestication. In this context, genome-wide high density SNP maps representing complex QTLs associated with critical agronomic traits can be used in the early stages of domestication and enables selection of inbred plants which maximizes optimal allelic segregation and fixing of important QTLs.

In principle, the methods according to the present invention can be applied to any plant. Therefore, monocotyledonous as well as dicotyledonous plant species are particularly suitable. The process is preferably used with plants that are important or interesting for agriculture, horticulture, for the production of biomass used in producing liquid fuel molecules and other chemicals, and/or forestry.

Thus, the invention has use over a broad range of plants, preferably higher plants pertaining to the classes of Angiospermae and Gymnospermae. Plants of the subclasses of the Dicotyledenae and the Monocotyledonae are particularly suitable. Monocotyledonous plants belong to the orders of the Magnoliidae, Illiciales, Laurales, Piperales Aristochiales, Nymphaeales, Ranunculales, Papaverales, Sarraceniales, Trochodendrales, Hamamelidales, Eucomiales, Leitneriales, Myricales, Fagales, Casuarinales, Caryophyllales, Batales, Polygonales, Plumbaginales, Dilleniales, Theales, Malvales, Urticales, Lycidiales, Violates, Salicales, Capparales, Ericales, Diapensiales, Ebenales, Primulales, Rosales, Fabales, Podostemales, Haloragales, Myrtales, Cornales, Proteales, Santales, Rafflesiales, Celastrales, Euphorbiales, Rhamnales, Sapindales, Juglandales, Geraniales, Polygalales, Umbellales, Gentianales, Polemoniales, Lamiales, Plantaginales, Scrophulariales, Campanulales, Rubiales, Dipsacales, and Asterales. Monocotyledonous plants belong to the orders of the Alismatales, Hydrocharitales, Najadales, Triuridales, Commelinales, Eriocaulales, Restionales, Poales, Juncales, Cyperales, Typhales, Bromeliales, Zingiberales, Arecales, Cyclanthales, Pandanales, Arales, Lilliales, and Orchidales. Plants belonging to the class of the Gymnospermae are Pinales, Ginkgoales, Cycadales and Gnetales.

Suitable species may include members of the genus Abelmoschus, Abies, Acer, Agrostis, Allium, Alstroemeria, Ananas, Andrographis, Andropogon, Artemisia, Arundo, Atropa, Berberis, Beta, Bixa, Brassica, Calendula, Camellia, Camptotheca, Cannabis, Capsicum, Carthamus, Catharanthus, Cephalotaxus, Chrysanthemum, Cinchona, Citrullus,
autumnale, Veratrum californica, Digitalis lanata, Digitalis purpurea, Dioscorea spp.,
Andrographis paniculata, Atropa belladonna, Datura stomonium, Berberis spp., Cephalotaxus
spp., Ephedra sinica, Ephedra spp., Erythroxyllum coca, Galanthus w workout, Scopolia spp.,
Lycopodium serratum (Huperzia serrata), Lycopodium spp., Rauwolfia serpentina, Rauwolfia
spp., Sanguinaria canadensis, Hyoscyamus spp., Calendula officinalis, Chrysanthemum
parthenium, Coleus forskohlii, Tanacetum parthenium, Parthenium argentatum (guayule),
Hevea spp. (rubber), Mentha spicata (mint), Mentha piperita (mint), Bixa orellana,
Alstroemeria spp., Rosa spp. (rose), Dianthus caryophyllus (carnation), Petunia spp. (petunia),
Poinsettia spp., (poinsettia), Nicotiana tabacum (tobacco), Lupinus albus (lupin),
Uniola paniculata (oats), bentgrass (Agrostis spp.), Populus tremuloides (aspen), Pinus spp.
(pine), Abies spp. (fir), Acer spp. (maple), Hordeum vulgare (barley), Poa pratensis
(bluegrass), Lolium spp. (ryegrass), Phleum pratense (timothy), and conifers. Of interest are
plants grown for energy production, so called energy crops, such as cellulose-based energy
crops like Panicum virgatum (switchgrass), Sorghum bicolor (sorghum, sudangrass),
Miscanthus giganteus (miscanthus), Saccharum sp. (energy cane), Populus balsamifera
(poplar), Andropogon gerardii (big bluestem), Pennisetum purpureum (elephant grass),
Phalaris arundinacea (reed canarygrass), Cynodon dactylon (bermudagrass), Festuca
arundinacea (tall fescue), Spartina pectinata (prairie cord-grass), Medicago sativa (alfalfa),
Arundo donax (giant reed), Secale cereale (rye), Salix spp. (willow), Eucalyptus spp.
(eucalyptus), Triticosecale spp. (triticum-wheat X rye), and Bamboo; and starch-based energy
crops like Zea mays (corn) and Manihot esculenta (cassava); and sucrose-based energy crops
like Saccharum sp. (sugarcane) and Beta vulgaris (sugar beet); and biofuel-producing energy
crops like Glycine max (soybean), Brassica napus (canola), Helianthus annuus (sunflower),
Carthamus tinctorius (safflower), Jatropha curcas (Jatropha), Ricinus communis (castor),
Elaeis guineensis (African oil palm), Elaeis oleifera (American oil palm), Cocos nucifera
(coconut), Camelina sativa (wild flax), Pongamia pinnata (Pongam), Olea europaea (olive),
Linum usitatissimum (flax), Crambe abyssinica (Abyssinian kale), and Brassica juncea.

The discussion of the general methods given herein is intended for illustrative purposes
only. Other alternative methods and embodiments will be apparent to those skilled in the art
upon review of this disclosure. The following examples are offered to illustrate, but not limit,
the invention.

A number of embodiments of the invention have been described. Nevertheless, it will
be understood that elements of the embodiments described herein can be combined to make
additional embodiments and various modifications may be made without departing from the
spirit and scope of the invention. Accordingly, other embodiments, alternatives and equivalents are within the scope of the invention and claimed herein. Headings within the applications are solely for the convenience of the reader, and do not limit in any way the scope of the invention or its embodiments.

All publications and patent applications mentioned in this specification are herein incorporated by reference to the same extent as if each individual publication or patent application was specifically can individually indicated to be incorporated by reference. The following examples are provided to illustrate various aspects of the present invention, and should not be construed as limiting the invention only to these particularly disclosed embodiments. The materials and methods employed in the examples below are for illustrative purposes, and are not intended to limit the practice of the present invention thereto. Any materials and methods similar or equivalent to those described herein can be used in the practice or testing of the present invention.

**EXAMPLES**

*Experimental design and results*

Most of the *Jatropha* parental lines currently used in production of *Jatropha* hybrids are heterozygous at many loci and thus lack complete genetic uniformity. As such, a *Jatropha* hybrid sexual crosses between two genetically dissimilar parents typically results in F1 hybrid offspring consisting of a combinatorial set of individual plants, each of which represents a specific and unique allelic segregation event and unique allelic combinations. Applicants have observed not only genotypic segregation among F1 individuals, but also dramatic phenotypic differences. For example, the seed yield from a productive cross can vary between individual plants by over 7 fold (FIG. 1).

When a large number of F1 hybrids from a single parental cross were planted and phenotypically characterized, Applicants observed a binomial yield distribution with most F1 plants exhibiting lower fruit yields. On the other hand, there often exists a small percentage of "transgressive segregant" plants containing specific sets of alleles which result in high yield productivity. This phenomenon is presumed to be a consequence of the heterozygosity of the parental lines which has resulted in each hybrid plant capturing a unique allelic segregation event among those loci that are heterozygous in one or the other or possibly both parents (FIG. 1). Applicants contemplate leveraging this phenomenon of combinatorial allelic segregation, much in the same way combinatorial chemistry is utilized for drug discovery. Recognizing that each hybrid cross can generate hundreds of individual allelic segregation events, one can
first identify the phenotypic outliers, for example the top 2-10% high seed yielding plants and the top 2-10% low seed yielding plants (FIG. 2). In some preferred embodiments, the top 5% high seed yielding plants and the top 5% low seed yielding plants are selected. In some particularly preferred embodiments, the top 2% high seed yielding plants and the top 2% low seed yielding plants are selected.

Thus in one aspect, Applicants contemplate utilizing genome-wide survey of SNPs (SNP fingerprints) to assist in identifying specific alleles associated with high yielding hybrid individuals that should be captured during parental line inbreeding. DNA is extracted from the tissues of the outlying plants and SNPs discovery can be performed either by genotyping by sequencing (GBS) or more preferably via whole genome re-sequencing and in-silico assembly. Subsequently, consensus 'high yield' and 'low yield' SNP fingerprints can be identified by applying statistical analysis and bioinformatics algorithms to the SNP fingerprints of the top performing individuals and bottom performing individuals that are derived from regionally productive hybrid. The designation of 'high yield' marker refers to a targeted SNP marker or markers that are found associated with high performing plants, and not found in association with the low yielding plants from the same cross. Similarly, a consensus 'low yielding' SNP fingerprint refers to a SNP marker or markers that are found associated with low yielding plants and not high yielding plants from the same cross. Next, these productive SNP combinations can be mapped back to the parental lines by a process in which the inheritance pattern of the 'high-yielding SNP fingerprints in the offspring are tracked back to the male and female parental lines by comparing the parental SNP fingerprints to those of their high-yielding offspring (FIG. 3). This process can shed light on and provide insight toward the precise segregation and recombination events taking place between the parental lines which resulted in the productive transgressive segregants. Such comparisons allows for identifying targeted SNPs that, when found in any one of the parental lines, have an increased likelihood of resulting in high-performing offspring. These results can then be used to select inbred offspring of the foregoing parental lines to accelerate the domestication (inbreeding) process.

As described in FIG. 4, in a next phase of the breeding method of the invention, the parental lines undergo inbreeding by either performing self-pollinations, or using tissue culture, or di-haploid methods. Several hundred plants of the inbreds are then planted in a green house, leaf tissues are collected, and genomic DNA is isolated. The genomic DNA from each of the inbred plants is subsequently subjected to SNP discovery and individual plant specific genome wide SNP maps are then compared to the consensus SNP maps associated with either high-yielding and low-yielding plants. Selection of the inbred plants to be taken forward in the
breeding process is then based on inbreds exhibiting segregation patterns that closely mimic predictions from which set of parental alleles contributed to the productive outliers. Since a number of these inbreds can be selected and crossed again to test for actual productivity gain, there exists a phenotypic feedback loop where the algorithms can be iteratively improved, and improvements can be made on predictability.

As it will be appreciated by one skilled in the art, the breeding methodologies disclosed herein have far reaching ramifications not only for utility in domestication of *Jatropha* but for other wild plants and animals. In addition, it is also contemplated that the methodologies of the invention have important applications in plant re-domestication. This is because in previous plant domestications, to the best of Applicants' knowledge, genome wide SNP mapping technologies have never been deployed to assess allelic segregation patterns during the first few cycles of inbreeding, the technologies simply did not exist. For example, by applying the methodologies of the invention in tomato domestication breeding programs, one could have potentially not only locked in for example, shelf life and ripening traits into tomatoes, but also flavor related attributes.

*Current Breeding workflow*

The workflow of a breeding program in accordance with the present invention typically involves the followings phases.

**Phase 1: Generation of Jatropha reference genomes.**

Applicants have generated a high-quality *Jatropha* reference genome by using a high-throughput sequencing method comprising the following steps

1) All-PathLG assembly with Illumina paired-end and mate-paired libraries.
2) Long read scaffolding of contigs using Ion Torrent. First-draft assembly was generated based on SOLiDs and 454 sequencing data.
3) Gene modeling and annotation.
4) Identification of gene models including compilation of the exome and regulome sequences.

**Phase 2: Evaluation and optimization of strategies for generating genome wide SNP fingerprints.**

This phase begins with the generation of simulated datasets of parental lines and their hybrid offspring with multiple parameters that can be varied - SNP density, yield, and offspring sample size. This is followed by a genome-wide statistical association analysis of simulated datasets, and the identification of the power given different offspring sample sizes and the ability to detect and association between a given fingerprint and yield. This phase typically involves the following steps:
1) Whole genome re-sequencing by using SOLiDs vs Ion Torrent vs Illumina sequencing platforms.
2) RNA-Seq
3) Genotyping by sequencing
4) Capturing Exome/regulome.

**Phase 3: Experimental validation**

Sexual crosses have been performed between two top performing parents originated from one geographical region (2000 plants per cross). Genome wide fingerprint maps are generated for approximately 300 individual plants, including individual parental lines and 50-150 top and 50-150 bottom yielding F1 offspring from each cross. Consensus high yield and consensus low yield SNP fingerprints have been generated using bioinformatics algorithms. Once generated, the algorithms for scoring SNPs in parental lines could be further refined and subsequently use to determine high yield SNP fingerprint for each parental line. Genomic DNA samples extracted from inbred progeny of the parental lines are then screened to determine which inbred offspring contained the targeted high yield SNP fingerprints. Approximately 100 inbred offspring from each selfed parental line are then planted and grown until flowering and then evaluated for various agronomic characteristics including seed yield and plant yield. Top performing hybrids are then reconstructed using the selected inbred offspring. Based on data training datasets of the newly acquired yield results, algorithms can be further refined.

In a scaled-up planting effort, approximately 2000 seedlings from top two hybrid crosses are planted. SNP fingerprint analyses are then performed on 50-150 each top and bottom yielding plants together with respective parental lines to generate consensus high yield and consensus low yield genome wide SNP maps. Based on data training datasets of the newly acquired yield results, algorithms can be further refined for scoring SNPs in parental lines. Refined algorithms are then used to determine preferred high yield SNP fingerprint for each of the parental lines. Genomic DNA samples extracted from seedlings of parental lines are then screened to determine which individual plants contain consensus high yielding SNP fingerprints. Parental individuals selected in this manner are identified as a target for inbreeding.

While a number of exemplary aspects and embodiments have been discussed above, those of skill in the art will recognize certain modifications, permutations, additions and sub-combinations thereof. It will be understood that elements of the embodiments described herein can be combined to make additional embodiments and various modifications may be made
without departing from the spirit and scope of the invention. Accordingly, other embodiments, alternatives and equivalents are within the scope of the invention and claimed herein. Headings within the applications are solely for the convenience of the reader, and do not limit in any way the scope of the invention or its embodiments. It is therefore intended that the following appended claims and claims hereafter introduced are interpreted to include all such modifications, permutations, additions and sub-combinations as are within their true spirit and scope of the invention or its embodiments.

The discussion of the general methods given herein is intended for illustrative purposes only. Other alternative methods and embodiments will be apparent to those skilled in the art upon review of this disclosure. It should also be understood that the examples provided herein are offered to illustrate, but not limit, the invention.
What I claimed is:

1. A method for selecting essentially homozygous parental plant lines from genetically-divergent, heterozygous plant lines for the production of hybrid progeny, comprising identifying one or more regions of the genomes of said essentially homozygous parental plant lines that, when present together in said hybrid progeny, result in improvement of a target trait.

2. The method of claim 1, wherein said identifying comprises the steps of:
   a) crossing a pair of genetically-divergent, heterozygous plant lines to create an offspring population segregating for said target trait;
   b) selecting from said segregating offspring population of step (a) a first offspring subpopulation and a second offspring subpopulation, wherein said first offspring subpopulation comprises a plurality of positive transgressive segregants, and said second offspring subpopulation comprises a plurality of negative transgressive segregants;
   c) identifying in the genomes of said segregants one or more regions of said genomes that are essentially associated with said first offspring subpopulation, but that are essentially absent in said second offspring subpopulation;
   d) selecting one or more genetically-divergent, heterozygous plant lines comprising genomic regions that are identical or homologous to said one or more regions of said genomes identified in step (c); and
   e) inbreeding said one or more genetically divergent, heterozygous plant lines selected in step (d) to produce essentially homozygous parental plant lines in which said homologous genomic regions have been fixed by said inbreeding.

3. The method of claim 1 or 2, wherein said one or more regions of said genomes are genetically associated with at least one genetic molecular marker.

4. The method of claim 3, wherein said at least one genetic molecular marker is selected from the group consisting of a SNP marker, an AFLP marker, a DAF marker, a RAPD marker, an FLP marker, and an SSR marker.

5. The method of any one of claims 2-4, wherein said essentially homozygous parental plant lines produced in step (e) are homozygous for said one or more regions of
6. The method of any one of claims 1-5, wherein said genetically-divergent, heterozygous plant lines are monocotyledonous or dicotyledonous plant lines.

7. The method of claim 6, wherein said genetically-divergent, heterozygous plant lines are lines of a domesticated or undomesticated crop.

8. The method of claim 6 or 7, wherein said genetically-divergent, heterozygous plant lines are *Jatropha* lines, maize lines, rice lines, sorghum lines, *Camelina* lines, *Pongamia* lines, *Brassica carinata* lines, castor lines, or tomato lines.

9. The method of claim 8, wherein said genetically-divergent, heterozygous plant lines are *Jatropha* lines.

10. The method of claim 9, wherein said *Jatropha* lines are selected from the group consisting of undomesticated *Jatropha*, inbred *Jatropha*, single cross F1 hybrid *Jatropha*, clonally propagated *Jatropha* clones, and a combination of any of the foregoing.

11. The method of any one of claims 1-10, wherein said target trait is selected from the group consisting of high productivity, synchronous flowering, high oilseed yield, day length insensitivity, growth habit uniformity, abiotic stress tolerance, biotic stress tolerance, water use efficiency, nitrogen use efficiency, and combinations of any thereof.

12. The method of claim 11, wherein said target trait is high productivity.

13. An essentially homozygous parental plant line produced by a method according to any one of claims 1-12.

14. The essentially homozygous parental plant line of claim 13, which is a *Jatropha* parental line.

15. A plant of said essentially homozygous parental plant line of claim 13 or 14.


17. The part of said plant of claim 16, which is selected from the group consisting of a protoplast, a cell, a tissue, an organ, a cutting, and an explant.

18. The part of said plant of claim 16, which is selected from the group consisting of an inflorescence, a flower, a sepal, a petal, a pistil, a stigma, a style, an ovary, an ovule, an embryo, a receptacle, a seed, a fruit, a stamen, a filament, an anther, a male or female gametophyte, a pollen grain, a meristem, a terminal bud, an axillary bud, a leaf, a stem, a root, a tuberous root, a rhizome, a tuber, a stolon, a corm, a bulb, an offset, a cell of
said plant in culture, a tissue of said plant in culture, an organ of said plant in culture, and a callus.

19. Progeny or seed of said plant of claim 15.

20. A method of producing a hybrid *Jatropha* seed, comprising crossing a first *Jatropha* plant of an essentially homozygous *Jatropha* parental line of claim 14 with a second *Jatropha* plant, and producing said hybrid *Jatropha* seed.

21. The method of claim 20, wherein said second *Jatropha* plant is of an essentially homozygous *Jatropha* parental line.

22. The method of claim 21, wherein said second essentially homozygous *Jatropha* parental line is produced by the method of any one of claims 1-12.

23. A hybrid *Jatropha* seed produced by the method of any one of claims 20-22.


25. A hybrid *Jatropha* plant according to claim 24, wherein said hybrid *Jatropha* plant exhibits an improved target trait.

26. A hybrid *Jatropha* plant according to claim 25, wherein said improved target trait is selected from the group consisting of high productivity, synchronous flowering, high oilseed yield, day length insensitivity, biotic stress tolerance, abiotic stress tolerance, water use efficiency, nitrogen use efficiency, growth habit uniformity, and combinations thereof.

27. A hybrid *Jatropha* plant according to claim 25 or 26, wherein said improved target trait is high productivity.

28. A part of said hybrid *Jatropha* plant of any one of claims 24-27.

29. The part of said hybrid *Jatropha* plant of claim 28, which is selected from the group consisting of a protoplast, a cell, a tissue, an organ, a cutting, and an explant.

30. The part of said hybrid *Jatropha* plant of claim 28, which is selected from the group consisting of an inflorescence, a flower, a sepal, a petal, a pistil, a stigma, a style, an ovary, an ovule, an embryo, a receptacle, a seed, a fruit, a stamen, a filament, an anther, a male or female gametophyte, a pollen grain, a meristem, a terminal bud, an axillary bud, a leaf, a stem, a root, a tuberous root, a rhizome, a tuber, a stolon, a corm, a bulb, an offset, a cell of said plant in culture, a tissue of said plant in culture, an organ of said plant in culture, and a callus.

31. Progeny or seed of said hybrid *Jatropha* plant of any one of claims 24-27.

32. Oil obtained from said seed of claim 31.
33. A product produced from said oil of claim 32, wherein said product is selected from the group consisting of biodiesel, biokerosene, hydraulic fluids, dielectric coolants, bioplastics, specialty chemicals, and pharmaceutical intermediates.
FIG. 1
FIG. 3