INVITED TECHNICAL REVIEW
Sunflower genetic, genomic and ecological resources

NOLAN C. KANE,*† JOHN M. BURKE,‡ LAURA MAREK,§ GERALD SEILER,¶ FELICITY VEAR,**
GREGORY BAUTE,† STEVEN J. KNAPP,†† PATRICK VINCOURT‡‡§§ and LOREN H. RIESEBERG†¶¶

*Department of Ecology and Evolutionary Biology, University of Colorado at Boulder, CB334, Ramaley N395, Boulder, CO 80309,
USA, †Department of Botany, University of British Columbia, 3529-6270 University Blvd., Vancouver, British Colombia, Canada V6T
1 Z4, ‡Department of Plant Biology, University of Georgia, 2502 Miller Plant Sciences, Athens, GA 30602, USA, §Oil Seed Crops,
North Central Regional Plant Introduction Station (NCRPIS), National Plant Germplasm System, Iowa State University, Ames, IA
50011, USA, ¶USDA, ARS, NCSL, 1307 18TH St N., Fargo, ND 58102-2765, USA, **INRA, UMR 1095, 234 Ave du Brezet, 63000
Clermont Ferrand, France, ††Monsanto Vegetable Seeds, 37437 State Hwy 16, Woodland, CA 95695, USA, ‡‡Laboratoire des
Interactions Plantes Micro-organismes, INRA, Chemin de Borde Rouge, BP 52627, 31326 Castanet Tolosan, France, §§Laboratoire des
Interactions Plantes Micro-organismes, CNRS, Chemin de Borde Rouge, BP 52627, 31326 Castanet Tolosan, France, ¶¶Department of
Biology, Indiana University, 1001 E. Third Street, Bloomington, IN 47405, USA

Abstract

Long a major focus of genetic research and breeding, sunflowers (*Helianthus*) are emerging as an increasingly important experimental system for ecological and evolutionary studies. Here, we review the various attributes of wild and domesticated sunflowers that make them valuable for ecological experimentation and describe the numerous publicly available resources that have enabled rapid advances in ecological and evolutionary genetics. Resources include seed collections available from germplasm centres at the USDA and INRA, genomic and EST sequences, mapping populations, genetic markers, genetic and physical maps and other forward- and reverse-genetic tools. We also discuss some of the key evolutionary, genetic and ecological questions being addressed in sunflowers, as well as gaps in our knowledge and promising areas for future research.

Keywords: adaptation, agriculture, angiosperms, ecological genetics, hybridization, speciation

Received 19 September 2011; revision received 22 August 2012; accepted 24 August 2012

Introduction

Iconic symbols in myth, art, politics and religion, sunflowers represent solar deities, power, nuclear nonproliferation, longevity and mortality. Over the past several decades, *Helianthus* has also emerged as an excellent experimental system for studying the ecological genetics of speciation, species boundaries, hybridization and domestication. With growing genomic resources, extensive publicly available seed collections, a rapidly developing genetic tool kit, important economic impacts and fascinating ecology, it is an ideal taxon for many ecological and evolutionary questions.

One of the core strengths of the system is the tremendous variation found within the genus. The diversity of speciation mechanisms and barriers to gene flow are truly remarkable, making it ideal for understanding speciation and divergence from many angles. The 49 named sunflower species, native to diverse habitats throughout most

of North America (Seiler & Rieseberg 1997), include examples of allo- and autopolyploids (Timme *et al.* 2007), ecologically isolated sympatric and allopatric species (Heiser *et al.* 1969), karyotypically divergent species (Chandler *et al.* 1986; Burke *et al.* 2004; Lai *et al.* 2005b), allopatric species with weak barriers to gene flow other than geography (Heiser *et al.* 1969) and several homoploid hybrid species (Rieseberg 1991). This variation has made *Helianthus* a model system for studying speciation (e.g. Rieseberg *et al.* 1995).

Key aspects of within- and among-species variation have been harnessed during domestication and improvement of several sunflower species, most prominently the Jerusalem artichoke (*H. tuberosus*), which was domesticated for its tuber, and the common sunflower *H. annuus*, cultivated worldwide for edible oil, edible seeds and the cut flower industry. Sunflower-breeding programmes have benefited from the introgression of wild germplasm from numerous annual species as well as from perennial species such as *H. tuberosus* and *H. giganteus* (Škorić 1992). Wild germplasm contains numerous ecologically important traits that can be useful in cultivation, including disease

Correspondence: Nolan C. Kane, Fax: 1 303 492 8699; E-mail: nckane@gmail.com

resistance (Liu *et al.* 2010), drought tolerance (Seiler *et al.* 2006) and cytoplasmic male sterility (CMS) and restoration (Jan 2000). New breeding efforts, of which many involve introgression of wild material, promise to substantially improve drought and salt tolerance in cultivated sunflowers and also may lead to cultivars that produce substantial cellulosic biofuel as a byproduct of cultivation for food. Also of economic importance are several weedy and invasive sunflowers (Muller *et al.* 2011), again including varieties of *H. annuus* as well as several other annual and perennial species that have escaped cultivation, been accidentally introduced to new habitats or evolved into noxious weeds from wild populations. Not only has weediness evolved in numerous sunflower species, novel weedy traits have also evolved within species multiple times (Kane & Rieseberg 2008; Muller *et al.* 2011). Herbicide tolerance is one such trait of *Helianthus* weeds, which makes them difficult to manage agricultural pest (Burton *et al.* 2004). Ironically, alleles conferring herbicide tolerance in wild sunflowers can be used to our advantage by introgressing them into the domesticated sunflower for weed control purposes. Clearly, for this and many other traits, a better understanding of basic ecology, evolution and genetics goes hand in hand with agricultural and other applied purposes, making sunflower an interesting system on numerous levels and broadening funding opportunities beyond those available for many ecologically and evolutionarily important wild species.

Here, we briefly review what is currently known about the phylogeny, natural history and genome evolution in *Helianthus;* we also describe the genetic and genomic resources available for the system and discuss several recent findings from analyses of the new genomic data sets.

Geographical range

Helianthus is indigenous to North America. Although the native ranges of most species are restricted to the continental United States, several extend into northern Mexico or southern Canada (Heiser *et al.* 1969). In addition to their indigenous distributions, numerous *Helianthus* species have become naturalized elsewhere in the world as a consequence of both intentional and inadvertent introductions by humans. Indeed, 22 *Helianthus* taxa are considered naturalized or invasive in Europe (Rehorek 1997), and sunflowers are abundant in parts of southern South America (Cantamutto *et al.* 2010) and southern and western Australia (Seiler *et al.* 2008).

Phylogeny

Because *Helianthus* is a recently evolved and species-rich group with a history of hybridization and polyploidy, phylogenetic reconstruction has proven to be challenging,

especially for the perennial species (Schilling & Heiser 1981; Schilling 1997; Schilling *et al.* 1998; Timme *et al.* 2007). Combining extensive crossability information (Heiser *et al.* 1969) with morphological characters for 49 species led to important insights, but crossability information is of limited utility in the perennials because of polyploidy (Schilling & Heiser 1981). The addition of genetic data from ribosomal genes clarified the relationships among most of the annual species, including the reticulation events forming the homoploid hybrid species (Rieseberg 1991). Chloroplast restriction fragments (Schilling 1997) and nuclear ribosomal internal transcribed spacer (ITS) sequence data (Schilling *et al.* 1998) shed light on relationships among the major lineages within the genus and also clarified the relationships between *Helianthus* and related genera (Schilling 2001), but failed to resolve finer-scale relationships among species due to paucity of informative characters. The phylogeny for *Helianthus* with the best resolution and support was generated by analysing sequence data from the external transcribed spacer of the 18S–25S nuclear ribosomal DNA region (Timme *et al.* 2007). This phylogeny successfully resolved relationships among the perennial species for the first time, identified the parentage of several hybrid and polyploid lineages and showed that an annual life history has evolved multiple times within the genus. Nonetheless, relationships among the basal lineages in the genus were poorly supported and discordant with phylogenetic analyses based on restriction site data from chloroplast DNA (Schilling 2001) and sequence data from the ITS of nuclear ribosomal DNA (Schilling *et al.* 1998).

To visualize the most well-understood relationships among species in the genus, we show relationships among five sections within the genus (Fig. 1a), as well as more detailed species relationships within the more widely studied annual clade section Helianthus (Fig. 1b). In Fig. 1a, we present a hypothetical phylogeny summarizing broad relationships of clades within the genus supported by multiple studies (Schilling & Heiser 1981; Rieseberg 1991; Schilling 1997, 2001; Schilling *et al.* 1998; Timme *et al.* 2007). The phylogeny presented in Fig. 1a differs from the traditional sectional classification of the genus, with sections Agrestis and Helianthus apparently monophyletic but sections Ciliares and Divaricati polyphyletic. Within section Ciliares, however, Series Ciliares (here Sect. Ciliares clade A) and Pumili (here Sect. Ciliares clade B) appear to be monophyletic. The series within Divaricati are more difficult to delineate, and the clades identified by molecular phylogenetic analyses do not correspond to any previous treatment. Species within each section are listed in Table 1, which also presents information on genome size, ploidy, life history and publicly available resources for each species as described in the next section.

Figure 1b presents a more detailed phylogeny for the annual section Helianthus, including the common sunflower *H. annuus,* the plains sunflower *H. petiolaris* and
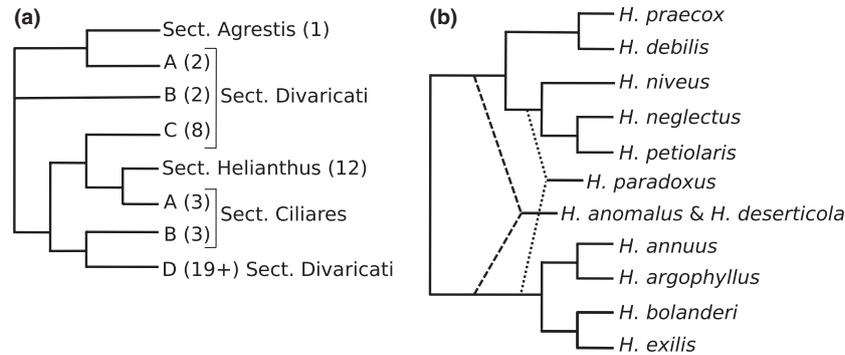
Fig. 1 Phylogenetic trees *for Helianthus*. (a) Phylogenetic tree for sections of the genus based on sequence analysis of the external transcribed spacer of nuclear ribosomal DNA (simplified from Timme *et al.* 2007). Numbers of species in each clade are given in parentheses following the section name. Note that sections *Ciliares* and *Divaricatus* are polyphyletic. (b) Phylogenetic network for section *Helianthus* based on inferences from nuclear ribosomal DNA analyses (Rieseberg 1991; Timme *et al.* 2007) and 11 single-copy nuclear genes (Moody & Rieseberg 2012). Putative hybrid speciation events are indicated by dashed lines.

their most closely related wild relatives. The two main lineages of annual sunflowers, those species related to *H. petiolaris* and those related to *H. annuus*, are well supported and distinct, although hybridization between the two lineages has given rise to at least three homoploid hybrid species, *H. paradoxus*, *H. anomalus* and *H. deserticola*. Relationships within the two lineages, however, are not uncontroversial, and some species delineations are not entirely clear. In particular, *H. debilis* is probably not monophyletic and may have to be split or combined with *H. praecox* (Timme *et al.* 2007), and the sister species pair *H. bolanderi* and *H. exilis* is often combined into a single species despite the strong differences in ecology and morphology (e.g. Heiser *et al.* 1969). A future goal should be to use the extensive EST data generated for *Helianthus* to develop a more accurate and highly supported phylogeny for the genus. Such a phylogeny could provide a framework for documenting the extent and timing of reticulate evolution in the genus, determining the genomic composition of hybrid and introgressed lineages, identifying sister taxa for studies of speciation and estimating the number and independence of ecological and evolutionary transitions within the genus.

Ecology and breeding system

Sunflowers occupy a broad range of habitats throughout their native and introduced ranges. The majority of species occupy open habitats such as disturbed areas, grasslands and deserts, but a few species do well in forest edges (*H. divaricatus* and *H. decapetalus*) or even grow as woodland understorey species (*H. microcephalus* and *H. radula*) and in seasonal bogs or marshes (*H. heterophyllus* and *H. paradoxus*). Several species are known for extreme abiotic stress tolerance, particularly the halophyte *H. paradoxus*, xeric species *H. deserticola*, dune-adapted species such as *H. anomalus* and *H. neglectus* and the serpentine soil

specialist *H. exilis*. Because multiple lineages of sunflowers have been introduced into non-native habitats, and because some species have been introduced into numerous locations independently, sunflowers are also emerging as a key model system for understanding the genomics and ecology of range expansions and the evolution of invasive taxa. Fascinating work has been done to document and explain invasive range expansion across the Argentine landscape (Cantamutto *et al.* 2010) and to understand the ecological conditions favouring invasion by particular *Helianthus* species (Cantamutto *et al.* 2008). Genetic and morphological work indicates that non-native populations have had multiple origins and that source populations probably include wild natives and crop-wild hybrids (Poverene & Cantamutto 2010; Muller *et al.* 2011). An emerging pattern is that range expansions in *Helianthus* are frequently associated with introgression from divergent lineages. This has been seen in the historical expansion of *H. annuus* into Texas, which involved introgression of herbivory tolerance and other traits (Heiser 1951; Whitney *et al.* 2006, 2010; Scascitelli *et al.* 2010) and California (Heiser 1949; Rieseberg *et al.* 1988; Carney *et al.* 2000), which involved high levels of interspecific hybridization and possibly adaptive introgression. Similarly, invasions onto other continents are known to have high levels of hybridization among species (Gutierrez *et al.* 2011) and between wild *H. annuus* and domesticated cultivars (Poverene & Cantamutto 2010; Muller *et al.* 2011; Lai *et al.* 2012). Whether this pattern is truly general, and whether this introgression is adaptive or due to demography or other neutral causes, remains a key question to be addressed in this system.

Most sunflowers are obligate outcrossers as a consequence of a sporophytic self-incompatibility system. Exceptions include *H. agrestis*, an annual species found in Florida, and all widely cultivated varieties of domesticated *H. annuus*, which are self-compatible (Heiser *et al.* 1969).

Table 1 Publicly available lines from the USDA and INRA, EST and genomic sequence data available from NCBI, characterized wild populations with information at the USDA, ploidy, chromosome numbers, genome size, life history and section for wild and domesticated sunflower species and hybrids. The 'habitat data' column lists the number of historical and current accessions for which habitat location and other data exist, regardless of whether the accession is still maintained. Letters following sections refer to the clades in Fig. 1

Taxon	USDA accessions	Habitat data	INRA accessions	NCBI nucleotide sequences	NCBI ESTs	Ploidy	Chromosome number	Genome size (Mb)	Life history	Section
Helianthus agrestis	9	11	1	5	0	2	34	12 691	Annual	Agrestis
H. angustifolius	22	28	1	5	0	2	34	5978	Perennial	Divaricati C
H. annuus cultivars	1867	NA	421	8189	10 5684	2	34	3528	Annual	Helianthus
H. annuus wild	930	1061	267	5908	28 014	2	34	3528	Annual	Helianthus
H. anomalus	6	14	3	194	0	2	34	5488	Annual	Helianthus
H. argophyllus	49	57	26	663	35 720	2	34	4336.5	Annual	Helianthus
H. arizonensis	2	5	2	5	0	2	34		Perennial	Ciliares B
H. atrorubens	14	21	2	7	0	2	34		Perennial	Divaricati D
H. bolanderi	7	14	5	4	0	2	34	4312	Annual	Helianthus
H. californicus	21	22	3	1	0	6	102		Perennial	Divaricati A
H. carnosus	2	3	0	3	0	2	34		Perennial	Divaricati C
H. ciliaris	26	27	2	7	21 590	4	68		Perennial	Ciliares B
H. cusickii	20	23	1	2	0	2	34		Perennial	Ciliares A
H. debilis	53	71	18	29	0	2	34	3577	Annual	Helianthus
H. decapetalus	30	32	4	6	0	2	34		Perennial	Divaricati D
H. deserticola	21	26	0	264	0	2	34	5635	Annual	Helianthus
H. divaricatus	26	45	3	5	0	2	34	8281	Perennial	Divaricati D
H. eggertii	12	15	2	7	0	6	102		Perennial	Divaricati C
H. exilis	30	35	2	4	33 961	2	34	4704	Annual	Helianthus
H. floridanus	8	13	1	2	0	2	34		Perennial	Divaricati C
H. giganteus	25	34	5	12	0	2	34	4728.5	Perennial	Divaricati D
H. glaucophyllus	11	13	2	3	0	2	34		Perennial	Divaricati D
H. gracilentus	6	11	1	2	0	2	34		Perennial	Ciliares A
H. grosseserratus	44	57	6	4	0	2	34		Perennial	Divaricati D
H. heterophyllus	17	26	0	8	0	2	34	4802	Perennial	Divaricati B
H. hirsutus	12	26	3	4	0	4	68		Perennial	Divaricati C
H. laciniatus	7	9	1	3	0	2	34		Perennial	Ciliares B
H. laevigatus	7	11	4	3	0	4	68		Perennial	Divaricati C
H. longifolius	3	4	0	3	0	2	34		Perennial	Divaricati D
H. maximilianii	64	91	23	7	0	2	34		Perennial	Divaricati D
H. microcephalus	13	21	2	4	0	2	34	4998	Perennial	Divaricati D
H. mollis	27	39	6	3	0	2	34		Perennial	Divaricati D
H. neglectus	28	29	3	132	0	2	34	3136	Annual	Helianthus
H. niveus	30	40	3	5	0	2	34	3577	Annual	Helianthus
H. nuttallii	42	46	20	4	0	2	34		Perennial	Divaricati D
H. occidentalis	15	21	2	4	0	2	34		Perennial	Divaricati D
H. paradoxus	2	10	2	268	30 517	2	34	5243	Annual	Helianthus
H. pauciflorus	46	59	3	3	0	6	102		Perennial	Divaricati D
H. petiolaris	139	185	23	1896	27 484	2	34	3332	Annual	Helianthus
H. porteri	8	9	1	7	0	2	34		Annual	Divaricati A
H. praecox	41	43	15	11	0	2	34	3454.5	Annual	Helianthus
H. pumilus	52	55	2	2	0	2	34		Perennial	Ciliares A
H. radula	37	47	1	3	0	2	34	5757.5	Perennial	Divaricati B
H. resinosus	23	31	4	3	0	6	102		Perennial	Divaricati D
H. salicifolius	19	24	3	5	0	2	34		Perennial	Divaricati D
H. schweinitzii	1	2	1	3	0	4	68		Perennial	Divaricati C
H. silphioides	15	21	2	2	0	2	34		Perennial	Divaricati D
H. simulans	4	4	1	7	0	2	34		Perennial	Divaricati C
H. smithii	6	10	2	1	0	2	34		Perennial	Divaricati D
H. strumosus	33	45	14	2	0	6	102		Perennial	Divaricati D
H. tuberosus	90	112	21	205	40 362	6	102	12 299	Perennial	Divaricati D

Table 1 (Continued)

Taxon	USDA accessions	Habitat data	INRA accessions	NCBI nucleotide sequences	NCBI ESTs	Ploidy	Chromosome number	Genome size (Mb)	Life history	Section
H. verticillatus	2	2	0	5	0	2	34		Perennial	Divaricati D
H. x laetiflorus	11	11	2	0	0	6	102		Perennial	Divaricati D
Helianthus hybrid	17	13	0	11	0	2	34		Annual	Various
All *Helianthus*	4052	2684	942	17 945	32 3332					

Both wild and domesticated sunflowers are pollinated by a diverse array of wild solitary bees, as well as domesticated honey bees (Neff & Simpson 1991; Sapir 2009). CMS segregates in the wild due to cytonuclear interactions (Serieys & Vincourt 1987; Rieseberg *et al.* 1994) and possibly plays a role in ecological adaptation and selection (Sambatti *et al.* 2008). In domesticated sunflowers, CMS was obtained for breeding and seed production purposes from hybridization between *H. petiolaris* and *H. annuus* by P. Leclercq (Leclercq 1969). Genes giving restoration of male fertility in the presence of this cytoplasm were obtained from both wild *H. annuus* (Kinman 1970) and the *H. petiolaris* accession which provided CMS (Leclercq 1971). Although most cultivated lines currently use CMS from this *H. petiolaris* material, over 60 more sources of CMS and 30 more sources of restorer alleles have been identified (see references in Jan 2000), suggestive of the large amount of diversity and the prevalence of CMS in wild *Helianthus*. As the specific genetic factors underlying CMS and restoration are identified and characterized, sunflower will probably become an important model for understanding cytonuclear interactions and their role in promoting or inhibiting speciation and interspecific gene flow. The advent of a reference nuclear and mitochondrial genome for this group as well as the many other resources being developed will surely promote this and related work.

Hybridization

Where species co-occur, hybridization leads to gene flow in many (but not all) cases, with widely varying outcomes. In some cases, gene flow is so low as to be virtually nonexistent, but in others it is quite high. *Helianthus annuus* ssp. *texanus* appears to have arisen through introgression from *H. debilis*, which apparently facilitated the southward expansion of *H. annuus* in Texas (Heiser 1951; Whitney *et al.* 2006, 2010). Rates of gene flow between these species are between four and seven migrants per generation (Scascitelli *et al.* 2010), but *H. debilis* remains strongly isolated and morphologically distinct from *H. annuus* and other sympatric species. In contrast, rampant hybridization has occurred between *H. annuus* and *H. bolanderi* since the introduction of *H. annuus* into California (Heiser 1949;

Rieseberg *et al.* 1988), with the result that many populations of *H. bolanderi* are predominantly hybrids. Interestingly, the closely related *H. exilis*, sister to *H. bolanderi* and also native to California, has not shared this fate, perhaps due to its strong ecological isolation (Carney *et al.* 2000): *H. exilis* thrives on serpentine soil, rarely coming into contact with *H. annuus* or other congeners. This suggests that the unique ecology of *H. exilis* may be a barrier to gene exchange, underlining the importance of ecology in promoting and maintaining isolation among diverging sunflower lineages.

Perhaps the best-studied hybridizing species are *H. annuus* and *H. petiolaris*, which co-occur over much of their broad ranges (Heiser 1947). Both species span the continental United States, from the Atlantic to Pacific and from Mexico to Canada. The primary ecological difference between the species appears to be that *H. annuus* prefers mesic clay soils while *H. petiolaris* thrives in drier sandy soils. These soil types often occur in close proximity, where the species can hybridize, resulting in low but measurable gene flow as a result, both at chloroplast (Dorado *et al.* 1992) and nuclear markers (Yatabe *et al.* 2007; Strasburg & Rieseberg 2008; Kane *et al.* 2009). These hybrid zones tend to be quite restricted, because the hybrids have extremely low levels of fertility (<1% viable seed), but because of the high effective population sizes of the two species, effective long-term levels of migration have been estimated as approximately 0.5 migrants per generation in each direction (Strasburg & Rieseberg 2008). Because of this low level of gene flow, divergence due to drift may be impeded, and advantageous or neutral alleles may move between the two species, but divergence will be high near loci where genetic background, ecology or other factors select against introgression.

Speciation

Evolutionary dead ends and low levels of gene flow are not the only results of hybridization. In several cases, hybridization between *H. annuus* and *H. petiolaris* has resulted in the origins of novel lineages of extremophiles (Rieseberg *et al.* 1990; Rieseberg 1991; Schwarzbach & Rieseberg 2002). Interestingly, these hybrid species (*H. anomalus*, *H. deserticola* and *H. paradoxus*) all evolved

without polyploidization, a process known as homoploid hybrid speciation. Because the origins of these species can in some ways be mimicked in the laboratory and the field by crossing the parental species, these speciation events have a particular fascination. Indeed, even without imposing ecological selection, crosses derived from *H. annuus* × *petiolaris* hybrids rapidly converge on a genetic make-up similar enough to the three natural homoploid hybrid sunflower species that they are largely interfertile with all three hybrid species and highly infertile with the parental species (Rieseberg 2000; Rieseberg *et al.* 2003; Lai *et al.* 2005b). Re-playing of evolutionary history is repeatable in only some respects: while hybrids from the laboratory and nature are interfertile and appear to have similar genetic make-up, their ecological niches are entirely different (Welch & Rieseberg 2002a,b; Rosenthal *et al.* 2005). Field experiments examining performance of synthetic hybrids in the homoploid hybrids' native environments have shown that the range of phenotypes in the hybrids exceeds that of the parents, to the extent that transgressive segregation may lead to some individuals with superior stress tolerance compared to either parent (Gross *et al.* 2004; Ludwig *et al.* 2004). However, because some aspects of the homoploids' genetic make-up and ecology were not recreated in these synthetic hybrids, it has been suggested that ecologically mediated selection throughout the process of homoploid hybrid speciation may be important (Karrenberg *et al.* 2007). Clearly, these homoploid hybrids are a system with many fascinating secrets yet to be revealed.

Adaptation and domestication

A new focus of research is on understanding phenotypic variation within and among sunflower species. This includes the study of local adaptation and the evolution of ecotypes within species, which may sometimes be the beginnings of a more long-term divergence (Kane & Rieseberg 2007). For instance, substantial variation in flowering phenology exists within *H. annuus* (Blackman *et al.* 2011a,b). Unlike other well-studied species, control of flowering time varies substantially in sunflowers, including short-day, long-day and day-neutral photoperiod responses (Blackman *et al.* 2011a,b). In the case of dune-adapted *H. petiolaris* in Colorado, differences in flowering time may be a key cause of reproductive isolation from nearby nondune populations (R. Andrew, personal communication).

Some of the most interesting within-species variation is in cultivated sunflower species, particularly *H. annuus*. The domesticated common sunflower has been cultivated for a variety of purposes and is highly diverse in terms of phenology and morphology. The genetics of sunflower domestication are becoming increasingly well understood (Burke *et al.* 2002, 2005; Baack *et al.* 2008), with a single cultivated lineage clearly originating in what is now the east–central United States (Harter *et al.* 2004; Wills & Burke

2006; Blackman *et al.* 2011a,b) and domestication having involved many genetic changes of small effect (Burke *et al.* 2002; Gandhi *et al.* 2005; Wills & Burke 2007). Some of the genes underlying important differences in phenotypes and phenology are also beginning to be understood, which has further supported our understanding of domestication and improvement (Blackman *et al.* 2010, 2011a,b). Additionally, characterization of the genes underlying iconic traits in mutated ornamental varieties has given insights into the genetics of floral patterning and inflorescence architecture and development (Chapman *et al.* 2012). Gene duplication and specialization of paralogous copies has been an important part of the evolution of inflorescence architecture (Chapman *et al.* 2008a, 2012; Thtiharju *et al.* 2012). As the genes underlying the regulation of development become increasingly well characterized, we will gain a better understanding of evolution of the compound inflorescence that defines the sunflower family (Asteraceae).

Genome evolution

Analysis of the age distribution of duplicate genes indicates that, while genetically diploid, *Helianthus* has experienced at least three episodes of polyploidy over the past 60 Myr (Barker *et al.* 2008; J. Bowers, personal communication). Two of these episodes appear to have occurred near the origin of the Compositae family, whereas the third duplication is coincident with an approximate doubling of chromosome numbers at the base of the sunflower tribe, Heliantheae. However, most of the genes that were duplicated during these ancient whole-genome duplications have been silenced or lost. As a consequence, the 3.6 Gb *H. annuus* genome is dominated by repetitive elements; a recent analysis of more than 100 Gb of whole-genome shotgun sequence indicates that 78.5% of the genome consists of repeats, mainly LTR retrotransposons (Kane *et al.* 2011).

There is interesting variation in genome size among *Helianthus* species with a fourfold size difference across the diploids. In general, the perennial species have larger genomes than most of the annuals (Sims & Price 1985; and see Table 1). Among the annual species of section *Helianthus*, there is a roughly 50% increase in genome size in the three homoploid hybrid species (*H. anomalus, H. deserticola, H. paradoxus*). Ungerer *et al.* (2006) showed that large-scale expansions of LTR retrotransposons are responsible for the genome size increases in the three hybrid lineages. Because this genome expansion occurred at least three times independently from hybridization from the same parental species, these lineages represent an ideal system to study what causes the activation and inactivation of retrotransposons, and what effects such expansions in retrotransposon copy number have on fitness, phenotypes and genome architecture. A baseline for this work has been set

by recent efforts to characterize in detail the repetitive content in *H. annuus* (Staton *et al.* 2012).

Helianthus is also well known for its high rate of karyotypic evolution, and closely related species typically differ by multiple chromosomal translocations and inversions (Heiser 1947; Chandler *et al.* 1986; Burke *et al.* 2004; Heesacker *et al.* 2009). The most striking example of rapid karyotypic change comes from comparative mapping studies of the three homoploid hybrid species and their parental species, *H. annuus* and *H. petiolaris* (Rieseberg *et al.* 1995; Lai *et al.* 2005b). Not only are the parental species highly divergent in karyotype, differing by at least eight translocations and three inversions, but the genomes of the three hybrid species are comprehensively restructured relative to their parents. Interestingly, most QTLs underlying hybrid pollen sterility in sunflowers map to chromosomal breakpoints (Lai *et al.* 2005a), and analyses of patterns of introgression across hybrid zones in nature suggest that interspecific gene flow is reduced in these genomic regions (Yatabe *et al.* 2007; Strasburg *et al.* 2009). However, marker densities on these maps are too low to detect microchromosomal rearrangements, and the location of chromosomal breakpoints cannot be inferred precisely. Thus, a more complete understanding of karyotypic evolution in *Helianthus* and its role in speciation will require high-density maps and genome-wide analyses of sequence divergence, both of which are underway. Because the annual sunflowers, in particular, co-occur in numerous distinct hybrid zones, and because these hybrid zones involve different species pairs in different locations, *Helianthus* represents an ideal model system for understanding the relationship between chromosomal evolution and speciation.

Germplasm and genomic resources

Supporting this extensive historical and ongoing research are numerous well-maintained, publicly available lines that can be obtained from germplasm centres in the United States (Marek *et al.* 2008, 2012) and elsewhere. A total of 2797 accessions of *H. annuus* and 1344 accessions of other species (Table 1), collected from native, cultivated and invasive populations from around the globe, are catalogued at the USDA's sunflower genebank in Ames, Iowa (http://www.ars-grin.gov/cgi-bin/npgs/html/site_holding. pl?NC7). A partially overlapping set of 521 wild lines and 421 cultivated accessions (including 331 old open-pollinated varieties and a core collection of 96 inbred lines) are also available from INRA's seed centres in France (Table 1, Table S1). In addition to making seed available, the USDA's website has an extensive database of measured traits, including morphology, phenology, chemistry (particularly seed oil characters) and genotypic information. These lines are thus both well maintained and well characterized, making it possible to identify subsets of

lines representing the majority of phenotypic and genetic variation within a species (Brothers & Miller 1999; Mandel *et al.* 2011).

In addition to inbred and open-pollinated accessions, several well-characterized mapping populations have been developed (e.g. Burke *et al.* 2002; Tang *et al.* 2002; Baack *et al.* 2008; Bowers *et al.* 2012). Crosses have been made between wild and domesticated lines (Burke *et al.* 2002; Wills & Burke 2007; Baack *et al.* 2008; Wills *et al.* 2010), elite cultivars and Native American landraces not subjected to modern breeding efforts (J.G. Barb *et al.*, personal communication), between elite modern breeding lines (Berry *et al.* 1995; Tang *et al.* 2002) and between species (Rieseberg *et al.* 1995; Burke *et al.* 2004; Lai *et al.* 2005b; Heesacker *et al.* 2009). The existing sunflower genetic maps can be viewed and compared using CMap at the Sunflower CMap database (http://www.sunflower. uga.edu/cmap/). A consensus map based on two RIL populations including 453 RIL has been also made available at http://www.heliagene.org/Web/public/mapping_ downy_mildew_resistance_genes.html. As of this writing, there are thus multiple genetic maps that contain many thousands of markers with known positions in the genomes of *H. annuus* (Bowers *et al.* 2012) and related species, including one of the highest-density genetic maps ever made, with 2.6 million mapped SNPs, for *H. annuus* (C. Grassa, personal communication). Currently being developed, a physical map containing unique sequence-based markers every 5–6 kb will be a major addition, greatly facilitating the ongoing assembly of the domesticated sunflower genome (Kane *et al.* 2011).

The current version of the *H. annuus* genome sequence represents over 80% of the sunflower genome and is comprised of tens of thousands of scaffolds (Kane *et al.* 2011). Incorporation of the physical and genetic maps, as well as substantial new long-read Roche 454 FLX Plus and GS FLX Titanium sequence data, is certain to improve these measures substantially. Substantial Illumina GAII whole-genome shotgun sequence data are publicly available (GenomeProject ID #64989) for *H. annuus*. In addition to the domesticated *H. annuus* genome, which has been the primary target of the sequencing effort, substantial genomic sequence data are being generated for numerous wild species, and high-quality annotated genomes will be assembled for several, including the annuals *H. argophyllus*, *H. petiolaris* and the perennial *H. maximiliani*.

Over 2000 microsatellites and other markers have been developed and genetically mapped (Berry *et al.* 1995; Gentzbittel *et al.* 1995, 1999; Jan *et al.* 1998; Gedil *et al.* 2001; Burke *et al.* 2002, 2004; Tang *et al.* 2002; Yu *et al.* 2003; Lai *et al.* 2005a,b; Chapman *et al.* 2008b; Heesacker *et al.* 2008, 2009), some of which occur in expressed genes and some in anonymous markers located largely in intergenic regions. Several large EST databases in sunflowers have

provided important additional resources for developing SSRs quickly and cost-effectively (Pashley *et al.* 2006; Ellis & Burke 2007) as have user-friendly bioinformatic pipelines for identifying potentially useful microsatellite markers (e.g. find SSR, Kane & Rieseberg 2007; Barker *et al.* 2010). These resources have greatly facilitated population genetic work characterizing diversity and gene flow within species (Liu & Burke 2005; Kane & Rieseberg 2007; Chapman *et al.* 2008b; Kane *et al.* 2009) and across species boundaries (Yatabe *et al.* 2007; Scascitelli *et al.* 2010). The origin of several of the homoploid species has also been investigated using these microsatellites (Schwarzbach & Rieseberg 2002; Welch & Rieseberg 2002b; Gross *et al.* 2003, 2007; Sapir *et al.* 2007).

With 323 332 Sanger ESTs available in GenBank as of April 2011, and millions of ESTs sequenced using next-generation technology (Lai *et al.* 2012), the genetic diversity at expressed genes is extremely well characterized in *Helianthus*. Sequence variation within species is quite high for both *H. annuus* and *H. petiolaris* (Strasburg *et al.* 2009). In addition to sequence data available on Genbank, information on the ESTs and other resources for sunflowers and related species can be found on the Compositae Genome Project (CGP) database (http://compgenomics.ucdavis. edu/) and at HeliaGene (http://www.heliagene.org/). Taking advantage of the available variation within *H. annuus*, two SNP chips have been constructed (Kane *et al.* 2011; Bachlava *et al.* 2012): a 10 640-SNP Infinium array and a 2.56 million feature Affymetrix chip, based on Sanger EST sequences from *H. annuus*, *H. petiolaris*, *H. argophyllus*, *H. exilis*, *H. paradoxus*, *H. ciliaris* and *H. tuberosus* sequenced by the CGP. Both have been used successfully for genetic mapping (Bowers *et al.* 2012; N. Langlade *et al.* unpublished; J. Bowers *et al.* unpublished), and high-density genetic maps are expected to be published using each chip in the near future. To assess expression differences throughout the transcriptome, both 4- and 12-plex NimbleGen expression arrays have been developed for common sunflower (Lai *et al.* 2012), with 68 400 features from 33 376 unigenes in the 4-plex array and 136 454 features from 48 683 unigenes from the 12-plex array. Additional gene expression work in progress uses a 44 000 feature Agilent array that is populated with genes from *H. argophyllus* and *H. annuus* (J. M. Burke, personal communication). Reliable genome size estimates have been made for 20 species, with chromosome numbers characterized for all species (Table 1). Replication for each species, more details on these estimates and the methods used can be found on the Kew c-values database (http://data.kew.org/cvalues/).

Reverse-genetic tools are less well developed than forward-genetic approaches in *Helianthus*. However, substantial work in this area is in progress, with tilling populations in Italy (Sabetta *et al.* 2009) and in France (P. Vincourt *et al.*, unpublished) and reports of successful transformation from several groups (Schrammeijer *et al.* 1990; Knittel *et al.* 1994; Malone-Shoneberg *et al.* 1994; Grayburn & Vick 1995; Rao & Rohini 1999; Weber *et al.* 2003; Mohamed *et al.* 2006; Manavella & Chan 2009; Zhang 2011). Nested core collections have been built from sunflower inbred lines of different origins, and recent work demonstrates that this is a powerful way to identify genetic variation underlying important segregating traits (Coque *et al.* 2008; Mandel *et al.* 2011).

Conclusions and future directions

With the forthcoming full genome sequence, new reverse-genetic tools, and increasingly well-characterized variation in important traits such as flowering time and herbivory, it is clear that sunflowers are emerging as an ecological model system with numerous advantages. Work towards more reliable transformation protocols and further development of mapping populations exploiting the substantial standing variation are also quite promising. As sequencing costs drop and genetic and germplasm resources continue to be improved, sunflower evolutionary geneticists are poised to answer key classic questions in ecology, evolution, development and population genetics. These include the following: What is the role of hybridization in adaptive evolution and speciation? What is the genetic basis of transgressive segregation? How do chromosomal differences get established and what is their role in reproductive isolation? How has the unique, compound inflorescence of the Asteraceae evolved, and what are the genetic factors underlying its development? What is the nature of species? What genetic differences define species boundaries? How does hybridization and gene flow between lineages affect adaptation and divergence? What is the genetic basis of parallel adaptations, such as invasiveness or adaptation to extreme environments? Particularly in the annual sunflowers, all of which are diploid, the wealth of population genetic, genomic and ecological data currently available and in progress means that many of these questions are on the verge of being answered.

Acknowledgements

References

Baack EJ, Sapir Y, Chapman MA, Burke JM, Rieseberg LH (2008) Selection on domestication traits and QTLs in crop-wild sunflower hybrids. *Molecular Ecology*, **17**, 666–677.

Bachlava E, Taylor CA, Tang S *et al.* (2012) SNP discovery and development of a high-density genotyping array for sunflower. *PLoS ONE*, **7**, e29814.

Barker MS, Kane NC, Kozik A *et al.* (2008) Multiple paleopolyploidizations during the evolution of the Asteraceae reveal parallel patterns of duplicate gene retention after millions of years. *Molecular Biology and Evolution*, **25**, 2445–2455.

Barker MS, Dlugosch KM, Dinh L *et al.* (2010) EvoPipes.net: bioinformatic pipelines and forums for ecological and evolutionary genomics. *Evolutionary Bioinformatics*, **6**, 143–149.

Berry ST, Leon AJ, Hanfrey CC *et al.* (1995) Molecular marker analysis of *Helianthus annuus* L.2. Construction of an RFLP linkage map for cultivated sunflower. *Theoretical and Applied Genetics*, **91**, 195–199.

Blackman BK, Strasburg JL, Michaels SD, Rieseberg LH (2010) The role of recently derived FT paralogs in sunflower domestication. *Current Biology*, **20**, 629–635.

Blackman BK, Michaels S, Rieseberg LH (2011a) Connecting the sun to flowering in sunflower adaptation. *Molecular Ecology*, **20**, 3503–3512.

Blackman BK, Scascitelli M, Kane NC *et al.* (2011b) Sunflower domestication alleles support single domestication center in eastern North America. *Proceedings of the National Academy of Sciences of the United States of America*, **108**, 14350–14365.

Bowers JE, Bachlava E, Brunick RL, Rieseberg LH, Knapp SJ, Burke JM (2012) Development of a 10,000 locus genetic map of the sunflower genome based on multiple crosses. *G3: Genes, Genomes, Genetics*, **2**, 721–729.

Brothers ME, Miller JF 1999. Core subset for the cultivated sunflower collection. Proceedings of the 21st Sunflower Research Workshop, National Sunflower Association, January 14–15, 1999, Fargo, ND, 21, 124–127.

Burke JM, Tang S, Knapp SJ, Rieseberg LH (2002) Genetic analysis of sunflower domestication. *Genetics*, **161**, 1257–1267.

Burke JM, Lai Z, Salmaso M *et al.* (2004) Comparative mapping and rapid karyotypic evolution in the genus *Helianthus*. *Genetics*, **167**, 449–457.

Burke JM, Knapp SJ, Rieseberg LH (2005) Genetic consequences of a recent selective sweep in cultivated sunflower. *Genetics*, **171**, 1933–1940.

Burton MG, Mortensen DA, Marx DB, Lindquist JL (2004) Factors affecting the realized niche of common sunflower. *Weed Science*, **52**, 779–787.

Cantamutto M, Poverene M, Peinemann N (2008) Multi-scale analysis of two annual *Helianthus* species naturalization in Argentina. *Agriculture, Ecosystems & Environment*, **123**, 69–74.

Cantamutto M, Presotto A, Fernandez MI, Alvarez D, Poverene M, Seiler G (2010) High infraspecific diversity of wild sunflowers (*Helianthus annuus* L.) naturally developed in central Argentina. *Flora*, **205**, 306–312.

Carney SE, Gardner KA, Rieseberg LH (2000) Evolutionary changes over the fifty year history of a hybrid population of sunflowers (*Helianthus*). *Evolution*, **54**, 462–474.

Chandler JM, Jan C, Beard BH (1986) Chromosomal differentiation among the annual *Helianthus* species. *Systematic Botany*, **11**, 354–371.

Chapman MA, Leebens-Mack JL, Burke JM (2008a) Positive selection and expression divergence following gene duplication in the sunflower *CYCLOIDEA* family. *Molecular Biology and Evolution*, **25**, 1260–1273.

Chapman MA, Pashley CH, Wenzler J *et al.* (2008b) A genomic scan for selection reveals candidates for genes involved in the evolution of cultivated sunflower (*Helianthus annuus*). *Plant Cell*, **20**, 2931–2945.

Chapman MA, Tang S, Draeger D *et al.* (2012) The genetics of floral symmetry in van Gogh's sunflowers reveals independent recruitment of CYCLOIDEA genes in the Asteraceae. *PloS Genetics*, **8**, 1–10.

Coque M, Mesnidrey S, Romestant M *et al.* (2008) Sunflower nested core collections for association studies and phenomics. Proceedings of the 17th International Sunflower Conference, June 8–12, 2008, Cordoba, Spain, 17, 725–728.

Dorado O, Rieseberg LH, Arias D (1992) Chloroplast DNA introgression in southern California sunflowers. *Evolution*, **46**, 566–572.

Ellis JR, Burke JM (2007) EST-SSRs as a resource for population genetic analyses. *Heredity*, **99**, 125–132.

Gandhi SD, Heesacker AF, Freeman CA, Argyris J, Bradford K, Knapp SJ (2005) The self-incompatibility locus (S) and quantitative trait loci for self-pollination and seed dormancy in sunflower. *Theoretical and Applied Genetics*, **111**, 619–629.

Gedil MA, Wye C, Berry S *et al.* (2001) An integrated restriction fragment length polymorphism – amplified fragment length polymorphism linkage map for cultivated sunflower. *Genome*, **44**, 213–221.

Gentzbittel L, Vear F, Zhang YX, Berville A, Nicolas P (1995) Development of a consensus linkage RFLP map of cultivated sunflower (*Helianthus annuus* L.). *Theoretical and Applied Genetics*, **90**, 1079–1086.

Gentzbittel L, Mestries E, Mouzeyar S *et al.* (1999) A composite map of expressed sequences and phenotypic traits of the sunflower (*Helianthus annuus* L.) genome. *Theoretical and Applied Genetics*, **99**, 218–234.

Grayburn WS, Vick BA (1995) Transformation of sunflower (*Helianthus annuus* L.) following wounding with glass beads. *Plant Cell Reports*, **14**, 285–289.

Gross BL, Schwarzbach AE, Rieseberg LH (2003) Origin(s) of the diploid hybrid species *Helianthus deserticola* (Asteraceae). *American Journal of Botany*, **90**, 1708–1719.

Gross BL, Kane NC, Lexer C *et al.* (2004) Reconstructing the origin(s) of *Helianthus deserticola*, survivorship and selection on the desert floor. *American Naturalist*, **164**, 145–156.

Gross BL, Turner KG, Rieseberg LH (2007) Selective sweeps in the homoploid hybrid species *Helianthus deserticola*, evolution in concert across populations and across origins. *Molecular Ecology*, **16**, 5246–5258.

Gutierrez A, Cantamutto M, Poverene M (2011) Persistence of sunflower crop traits and fitness in *Helianthus petiolaris* populations. *Plant Biology*, **13**, 821–830.

Harter AV, Gardner KA, Falush D, Lentz DL, Bye R, Rieseberg LH (2004) Origin of extant domesticated sunflowers in eastern North America. *Nature*, **430**, 201–205.

Heesacker A, Kishore VK, Gao W *et al.* (2008) SSRs and INDELs mined from the sunflower EST database, abundance, polymorphisms, and cross-taxa utility. *Theoretical and Applied Genetics*, **117**, 1021–1029.

Heesacker A, Bachlava E, Brunick RL, Burke JM, Rieseberg LH, Knapp SJ (2009) Chromosomal rearrangements and unbalanced segmental duplications underlying the evolution of the silverleaf and common sunflower genomes. *Plant Genome*, **2**, 233–246.

Heiser CB (1947) Hybridization between the sunflower species *Helianthus annuus* and *petiolaris*. *Evolution*, **1**, 249–262.

Heiser CB (1949) Studies in the evolution of the sunflower species *Helianthus annuus* and *H. bollanderi*. *University of California Publications in Botany*, **23**, 157–196.

Heiser CB (1951) Hybridization in the annual sunflowers, *Helianthus annuus* × *H. debilis* var. *cucumerifolius*. *Evolution*, **5**, 42–51.

Heiser CB, Smith DM, Clegenger S, Martin WC (1969) The North American sunflowers (*Helianthus*). *Memoirs of the Torrey Botanical Club*, **22**, 1–218.

Jan CC (2000) Cytoplasmic male sterility in two wild *Helianthus annuus* L. accessions and their restoration. *Crop Science*, **40**, 1535–1538.

Jan CC, Vick BA, Miller JF, Kahler AL, Butler ET (1998) Construction of an RFLP linkage map for cultivated sunflower. *Theoretical and Applied Genetics*, **96**, 15–22.

Kane NC, Rieseberg LH (2007) Selective sweeps reveal candidate genes for adaptation to drought and salt tolerance in common sunflower, *Helianthus annuus*. *Genetics*, **175**, 1803–1812.

Kane NC, Rieseberg LH (2008) Genetics and evolution of weedy *Helianthus annuus* populations, adaptation of an agricultural weed. *Molecular Ecology*, **17**, 384–394.

Kane NC, King MG, Barker MS *et al.* (2009) Comparative genomic and population genetic analyses indicate highly porous genomes and high levels of gene flow between divergent *Helianthus* species. *Evolution*, **63**, 2061–2075.

Kane NC, Gill N, King MG *et al.* (2011) Progress towards a reference genome for sunflower. *Botany*, **89**, 429–437.

Karrenberg S, Lexer C, Rieseberg LH (2007) Reconstructing the history of selection during homoploid hybrid speciation. *American Naturalist*, **169**, 725–737.

Kinman ML (1970) Letter to participants. Proceedings of the 4th International Sunflower Conference, Memphis, USA, 4, 181–183.

Knittel N, Gruber V, Hahne G, Lénée P (1994) Transformation of sunflower (*Helianthus annuus* L.), a reliable protocol. *Plant Cell Reports*, **14**, 81–86.

Lai Z, Livingstone K, Zou Y et al. (2005a) Identification and mapping of SNPs and candidate genes in sunflower, towards a functional map. *Theoretical and Applied Genetics*, **111**, 1532–1544.

Lai Z, Nakazato T, Salmaso M et al. (2005b) Extensive chromosomal repatterning and the evolution of sterility barriers in hybrid sunflower species. *Genetics*, **171**, 291–303.

Lai Z, Kane NC, Kozik A et al. (2012) Genomics of Compositae weeds, EST libraries, microarray, and evidence of introgression. *American Journal of Botany*, **99**, 209–218.

Leclercq P (1969) Une stérilité mâle cytoplasmique chez le tournesol. *Annales de l'Amélioration des Plantes*, **19**, 99–106.

Leclercq P (1971) La stérilité mâle cytoplasmique du tournesol 1. Premières études sur la restauration de la fertilité. *Annales de l'Amélioration des Plantes*, **21**, 45–54.

Liu A, Burke JM (2005) Patterns of nucleotide diversity in wild and cultivated sunflower. *Genetics*, **173**, 321–330.

Liu Z, Cai X, Seiler GJ, Gulya TJ, Rashid KY, Jan C (2010) Transferring sclerotinia resistance genes from wild *Helianthus* species into cultivated sunflower. Proceedings of the 32nd Sunflower Research Workshop, National Sunflower Association, January 10–11, 2008, Fargo, ND, pp. 1–5.

Ludwig F, Rosenthal D, Johnston JA et al. (2004) Selection on leaf ecophysiological traits in a hybrid *Helianthus* species and early generation hybrids in a desert dune habitat. *Evolution*, **58**, 2682–2692.

Malone-Shoneberg J, Scelonge CJ, Burrus M, Bidney DL (1994) Stable transformation of sunflower using *Agrobacterium* and split embryonic axis explants. *Plant Science*, **103**, 199–207.

Manavella PA, Chan RL (2009) Transient transformation of sunflower leaf discs via an *Agrobacterium*-mediated method, applications for gene expression and silencing studies. *Nature Protocols*, **4**, 1699–1707.

Mandel JL, Dechaine JM, Markek LF, Burke JM (2011) Genetic diversity and population structure in cultivated sunflower and a comparison to its wild progenitor, *Helianthus annuus* L. *Theoretical and Applied Genetics*, **123**, 693–704.

Marek LF, Block CC, Gardner C (2008) 2008 Update, The USDA sunflower collection at the North Central Regional Plant Introduction Station, Ames, IA, USA. Proceedings of the 17th International Sunflower Conference, Cordoba, Spain, 17, 735–740.

Marek LF, Block CC, Gardner C (2012) 2012 Update: New sunflower genetic resources in the US national sunflower collection and potential use for crop improvement. Proceedings of the 18th International Sunflower Conference, Mar del Plata, Argentina, 18, 663–668.

Mohamed S, Boehm R, Schnabl H (2006) Stable genetic transformation of high oleic *Helianthus annuus* L. genotypes with high efficiency. *Plant Science*, **5**, 546–554.

Moody ML, Rieseberg LH (2012) Sorting through the chaff, nDNA gene trees for phylogenetic inference and hybrid identification of annual sunflowers (*Helianthus* sect. Helianthus). *Molecular Phylogenetics and Evolution*, **64**, 145–155.

Muller M-H, Latreille M, Tollon C (2011) The origin and evolution of a recent agricultural weed, population genetic diversity of weedy populations of sunflower (*Helianthus annuus* L.) in Spain and France. *Evolutionary Applications*, **4**, 499–514.

Neff JL, Simpson BB (1991) Nest biology and mating behavior of *Megachile fortis* in central Texas (Hymenoptera, Megachilidae). *Journal of the Kansas Entomological Society*, **64**, 324–336.

Pashley CH, Ellis JR, McCauley DE, Burke JM (2006) EST databases as a source for molecular markers: lessons from *Helianthus*. *Journal of Heredity*, **97**, 381–388.

Poverene M, Cantamutto M (2010) A comparative study of invasive *Helianthus annuus* populations in their natural habitats of Argentina and Spain. *Helia*, **33**, 63–74.

Rao KS, Rohini VK (1999) *Agrobacterium*-mediated transformation of sunflower (*Helianthus annuus* L.), a simple protocol. *Annals of Botany*, **83**, 347–354.

Rehorek V (1997) Cultivated and escaped perennial *Helianthus* species in Europe. *Preslia*, **69**, 59–70.

Rieseberg LH (1991) Homoploid reticulate evolution in *Helianthus*, evidence from ribosomal genes. *American Journal of Botany*, **78**, 1218–1237.

Rieseberg LH (2000) Crossing relationships among ancient and experimental sunflower hybrid lineages. *Evolution*, **54**, 859–865.

Rieseberg LH, Soltis DE, Palmer JD (1988) A molecular reexamination of introgression between *Helianthus annuus* and *H. bolanderi*. *Evolution*, **42**, 227–238.

Rieseberg LH, Carter R, Zona S (1990) Molecular tests of the hypothesized hybrid origin of two diploid *Helianthus* species (Asteraceae). *Evolution*, **44**, 1498–1511.

Rieseberg LH, Van Fossen C, Arias D, Carter RL (1994) Cytoplasmic male sterility in sunflower, origin, inheritance, and frequency in natural populations. *Journal of Heredity*, **85**, 233–238.

Rieseberg LH, Van Fossen C, Desrochers A (1995) Hybrid speciation accompanied by genomic reorganization in wild sunflowers. *Nature*, **375**, 313–316.

Rieseberg LH, Raymond O, Rosenthal DM et al. (2003) Major ecological transitions in annual sunflowers facilitated by hybridization. *Science*, **301**, 1211–1216.

Rosenthal DM, Rieseberg LH, Donovan LA (2005) Recreating ancient hybrid species' complex multi-trait phenotypes from early generation synthetic hybrids, three examples using wild sunflowers. *American Naturalist*, **166**, 26–41.

Sabetta W, Montemurro C, Alba V, Alba E, Sunseri F, Blanco A (2009) SunTILL, A sunflower tilling platform for functional analysis of genes involved in fatty acid biosynthesis. Proceedings of the 53rd Italian Society of Agricultural Genetics Annual Congress, ISBN 978-88-900622-9-2. Poster 7.58.

Sambatti JMB, Ortiz-Barrientos D, Baack EJ, Rieseberg LH (2008) Ecological selection maintains cytonuclear incompatibilities in hybridizing sunflowers. *Ecology Letters*, **11**, 1082–1091.

Sapir Y (2009) Effects of floral traits and plant genetic composition on pollinator behavior. *Arthropod-Plant Interactions*, **3**, 115–129.

Sapir Y, Moody ML, Brouillette LC, Donovan LA, Rieseberg LH (2007) Patterns of genetic diversity and candidate genes for ecological divergence in a homoploid hybrid sunflower, *Helianthus anomalus*. *Molecular Ecology*, **16**, 5017–5029.

Scascitelli M, Whitney KD, Randell RA, King M, Buerkle CA, Rieseberg LH (2010) Genome scan of hybridizing sunflowers from Texas (*Helianthus annuus* and *H. debilis*) reveals asymmetric patterns of introgression and small islands of genomic differentiation. *Molecular Ecology*, **19**, 521–541.

Schilling EE (1997) Phylogenetic analysis of *Helianthus* (Asteraceae) based on chloroplast DNA restriction site data. *Theoretical and Applied Genetics*, **94**, 925–933.

Schilling EE (2001) Phylogeny of *Helianthus* and related genera. *Oléagineux, Corps Gras, Lipides*, **8**, 22–25.

Schilling EE, Heiser CB (1981) An infra-generic classification of *Helianthus* (Compositae). *Taxon*, **30**, 393–403.

Schilling EE, Linder CR, Noyes R, Rieseberg LH (1998) Phylogenetic relationships in *Helianthus* (Asteraceae) based on nuclear ribosomal DNA internal transcribed spacer region sequence data. *Systematic Botany*, **23**, 177–188.

Schrammeijer B, Sijmons PC, Elzen PJM, Hoekema A (1990) Meristem transformation of sunflower via *Agrobacterium*. *Plant Cell Reports*, **9**, 55–60.

Schwarzbach AE, Rieseberg LH (2002) Likely multiple origins of a diploid hybrid sunflower species. *Molecular Ecology*, **11**, 1703–1717.

Seiler GJ, Rieseberg LH (1997) Systematics, origin, and germplasm resources of the wild and domesticated sunflower. In: *Sunflower Science and Technology* (ed. Schneiter AA), pp. 21–66. ASA, CSSA, and ASSA, Madison, Wisconsin.

Seiler GJ, Gulya TJ, Marek LF (2006) Exploration for wild *Helianthus* species from the desert southwestern USA for potential drought tolerance. *Helia*, **29**, 1–10.

Seiler GJ, Gulya TJ, Kong G, Thompson S, Mitchell J (2008) Collection of wild naturalized sunflowers from the land down under. Proceedings of the 30th Sunflower Research Workshop, National Sunflower Association, January 10–11, 2008, Fargo, ND. Available from http://www.sunflowernsa.com/uploads/research/317/Seiler_etal_DownUnder_08.pd.

Serieys H, Vincourt P (1987) Characterisation of some new CMS sources from the *Helianthus* genus. *Helia*, **10**, 9–13.

Sims L, Price JH (1985) Nuclear DNA content variation in *Helianthus* (Asteraceae). *American Journal of Botany*, **72**, 1213–1219.

Škorić D, Atlagic J, Mainkovic R, Vasiljevic L, Mihaljcevic M, Masirevic S (1987) Genetic evaluation of Helianthus wild species and their use in breeding programs. FAO Sunflower Research Subnetwork, pp. 1–35. In: *Proc. 1984–1986 FAO Sunflower Subnetwork Progr. Rep.*, Szeged, Hungary. 28–30 July. FAO, Rome, Italy.

Staton SE, Bakken BH, Ungerer MC *et al.* (2012) The sunflower (*Helianthus annuus* L.) genome reflects a history of biased accumulation of transposable elements. *The Plant Journal*, doi:10.1111/j.1365-313X.2012.05072.x.

Strasburg JL, Rieseberg LH (2008) Molecular demographic history of the annual sunflowers *Helianthus annuus* and *H. petiolaris* – large effective population sizes and rates of long-term gene flow. *Evolution*, **62**, 1936–1950.

Strasburg JL, Scotti-Saintagne C, Scotti I, Lai Z, Rieseberg LH (2009) Genomic patterns of adaptive divergence between chromosomally differentiated sunflower species. *Molecular Biology and Evolution*, **26**, 1341–1355.

Tang S, Yu JK, Slabaugh MB, Shintani DK, Knapp SJ (2002) Simple sequence repeat map of the sunflower genome. *Theoretical and Applied Genetics*, **105**, 1124–1136.

Thtiharju S, Rijpkema AS, Vetterli A, Albert VA, Teeri TH, Elomaa P (2012) Evolution and diversification of the CYC/TB1 gene family in Asteraceae – a comparative study in *Gerbera* (Mutisieae) and sunflower (Heliantheae). *Molecular Biology and Evolution*, **29**, 1155–1166.

Timme RE, Simpson BB, Linder CR (2007) High-resolution phylogeny for *Helianthus* (Asteraceae) using the 18S-26S ribosomal DNA external transcribed spacer. *American Journal of Botany*, **94**, 1837–1852.

Ungerer MC, Strakosh S, Zhen Y (2006) Genome expansion in three hybrid sunflower species is associated with retrotransposon proliferation. *Current Biology*, **16**, r872–r873.

Weber S, Friedt W, Landes N *et al.* (2003) Improved *Agrobacterium* -mediated transformation of sunflower (*Helianthus annuus* L.), assessment of macerating enzymes and sonication. *Plant Cell Reports*, **21**, 475–482.

Welch ME, Rieseberg LH (2002a) Habitat divergence between a homoploid hybrid sunflower species, *Helianthus paradoxus* (Asteraceae), and its progenitors. *American Journal of Botany*, **89**, 472–478.

Welch ME, Rieseberg LH (2002b) Patterns of genetic variation suggest a single, ancient origin for the diploid hybrid species *Helianthus paradoxus*. *Evolution*, **56**, 2126–2137.

Whitney KD, Randell RA, Rieseberg LH (2006) Adaptive introgression of herbivore resistance traits in the weedy sunflower *Helianthus annuus*. *American Naturalist*, **167**, 794–807.

Whitney KD, Randell RA, Rieseberg LH (2010) Adaptive introgression of abiotic tolerance traits in the weedy sunflower *Helianthus annuus*. *New Phytologist*, **187**, 230–239.

Wills DM, Burke JM (2006) Chloroplast DNA variation confirms a single origin of domesticated sunflower (*Helianthus annuus* L.). *Journal of Heredity*, **97**, 403–408.

Wills DM, Burke JM (2007) Quantitative trait locus analysis of the early domestication of sunflower. *Genetics*, **176**, 2589–2599.

Wills DM, Abdel-Haleem H, Knapp SJ, Burke JM (2010) Genetic architecture of novel traits in the Hopi sunflower. *Journal of Heredity*, **101**, 727–736.

Yatabe Y, Kane NC, Scotti-Saintagne C, Rieseberg LH (2007) Rampant gene exchange across a strong reproductive barrier between the annual sunflowers, *Helianthus annuus* and *H. petiolaris*. *Genetics*, **175**, 1883–1893.

Yu JK, Tang S, Slabaugh MB *et al.* (2003) Towards a saturated molecular genetic linkage map for cultivated sunflower. *Crop Science*, **43**, 367–387.

Zhang Z (2011) Development of a transformation and regeneration system for sunflower (*Helianthus annus*). *In Vitro Cellular and Developmental Biology – Animal*, **47**(Suppl 1), 50–79.

All authors contributed to the writing of the paper and preparation of tables and figures.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Table S1 Domesicated cultivar populations available from the INRA seed centres.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.