Chapter 1

Australian Cotton Germplasm Resources

Warwick N. Stiller and Iain W. Wilson

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/58414

1. Introduction

1.1. History

Cotton has been sporadically cultivated in Australia since the late 1700s with varying degrees of success until the advent of the modern intensively managed and mechanised industry in the latter half of the Twentieth Century. There were initial attempts to establish an industry during the mid 1800s to fill the demand in Britain for lint caused by declining production during the US Civil War, but with the return of US production after the war this new incarnation of an Australian industry was relatively short lived. In the early 1900s, financial incentives were introduced by the Queensland State Government to promote production in that State and by 1934 around 4000 tonnes of cotton lint, primarily from imported cultivars, were being produced. However, by the mid 1900s production was again virtually non-existent because of the prohibitively high cost of labour [1]. By the 1960s, the development of proper infrastructure for irrigation and mechanised pickers led to a rapid increase in area and production in northern NSW and southern and central Qld which was aided by the development of locally adapted cultivars in the 1980s and 1990s. An industry was also attempted in the Ord River Irrigation Area in the tropical region of Western Australia. However, this had collapsed by 1973 predominantly due to insect resistance to pesticides, but the crop continued to flourish in the Eastern part of Australia where the insect pressure was lower. Genetically modified cotton (Ingard®-Monsanto's MON531 *Cry1Ac* Bt) was first commercialised in 1996 and by 2005 over 90% of the area was grown to GM cultivars (Monsanto's Bollgard II® and Roundup Ready Flex®). In the 2011/12 season, the cotton area exceeded 700,000 ha, was >99% GM and produced over one million tonnes of lint.

The Australian cotton industry covers a wide north-south geographical range of 13° latitude (Figure 1); is based on a combination of heavy clay soils; and is challenged by a range of pests and diseases. The industry is predominantly upland cotton (*Gossypium hirsutum*), with up to

2% of pima cotton (*Gossypium barbadense*). Since 2010, as a result of a long drought and eventually loss of infrastructure to support its processing, there has been no Pima cotton grown in this country. The Commonwealth Scientific and Industrial Research Organisation (CSIRO) cotton breeding program has developed and released over 100 very successful cultivars adapted to local conditions with a high yield and quality. By 2000, these had displaced all imported cultivars and any other local competition and currently all cotton grown commercially in Australia stems from that program. Between 1995 and 2009, yield progress from the CSIRO program was an impressive 18.3 kg ha^{-1} $year^{-1}$, with 48% of that yield improvement directly attributed to cultivar, 28% to improved management and 24% to the interaction between cultivar and management [2], making Australia one of the highest yielding cotton producers per hectare globally.

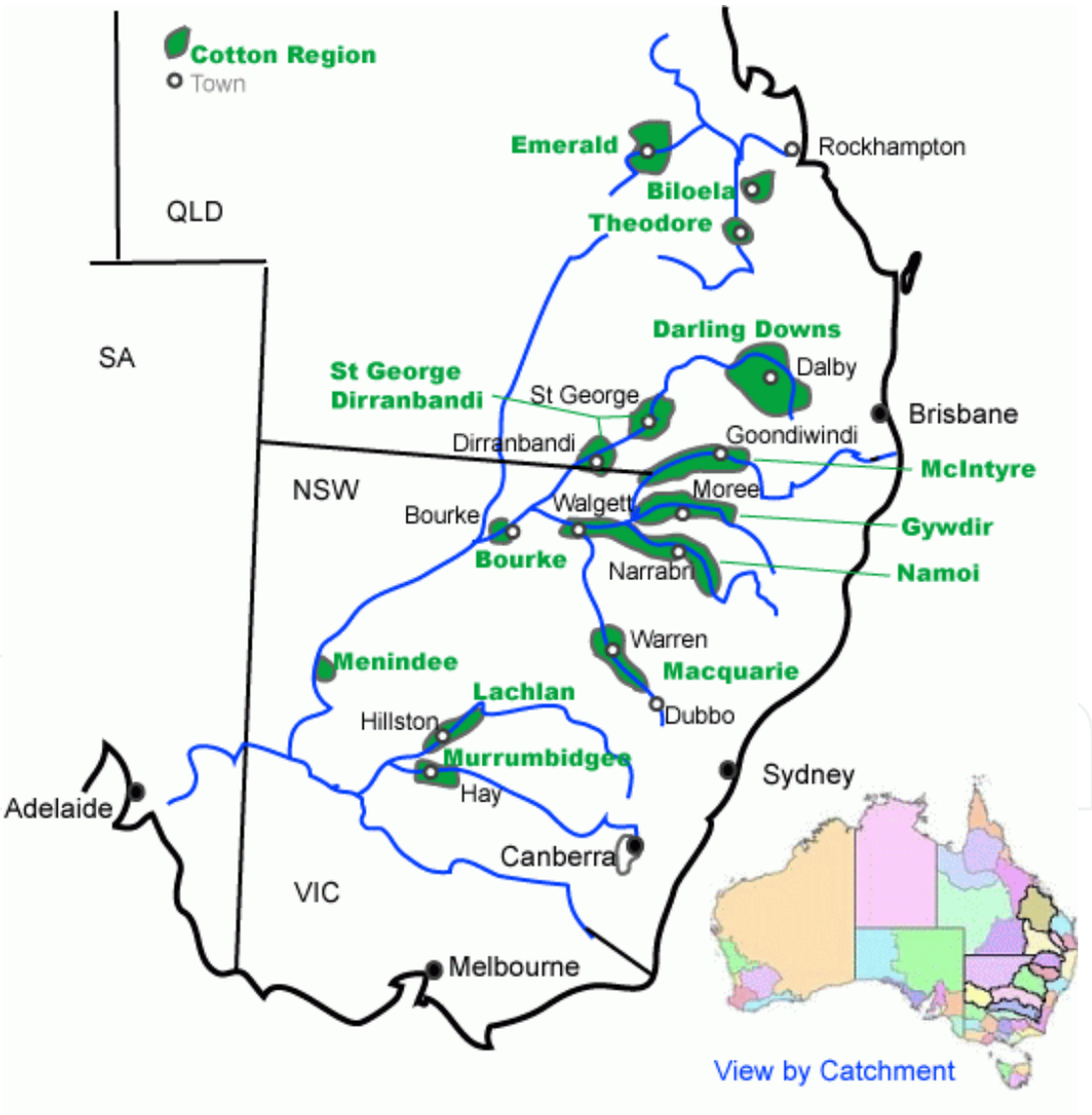

Figure 1. Australian cotton production regions

http://dx.doi.org/10.5772/58414

1.1.1. History of cotton breeding programs in Australia

Cotton breeding in Australia began in a very piecemeal manner with most early growers relying on unimproved and imported US cultivars. The first concerted effort to breed locally adapted cultivars came when a cotton research station was established in Central Queensland in 1923 and a Queensland DPI breeding program eventually released local cultivars to replace some of the imported cultivars. The first locally selected cultivar was Miller 43-9-0. This program also worked on insect host-plant resistance, producing a jassid-resistant hairy cultivar that was discontinued with the advent of widespread mechanical harvesting that was less tolerant of hairy leaf trash adhering to the cotton. This breeding program closed in 1990.

With bounties to encourage production and construction of dams to facilitate irrigation, cotton expanded into more temperate areas in the 1960s and particularly through the 1980s. In 1972, breeding programs in southern NSW (CSIRO) and in northern Western Australia (CSIRO) were all consolidated at one location at Narrabri in northern New South Wales (NSW), where NSW Agriculture had established a breeding program in 1960. The NSW Agriculture program initially screened overseas cultivars for suitability to the unique combination of soils and climate in northern NSW, but evolved to study resistance to Verticillium wilt. This program closed in 1980. The CSIRO breeding program in southern NSW had a particular emphasis on earliness and suitability for ultra narrow row systems. This program used a wide range of germplasm for breeding, including from USSR and US Texas material to produce a number of early-maturing cultivars such as Riverina Poplar and Riverina Gold [3]. This program closed in 1972 with the move to Narrabri.

Until the mid-1980s, the majority of cultivars grown in Australia were still imported from the US, principally Deltapine Smoothleaf, DP16, DP61 and DP90. The Delta and Pineland Company established a breeding program in 1990 and marketed a number of cultivars, particularly GM, over the next decade. This US company changed hands and withdrew from marketing cultivars in 2010 as CSIRO developed cultivars had displaced the majority of even their locally bred offerings, but it still retains a breeding station in Australia.

The CSIRO breeding program in northern Western Australia originally concentrated on adaptation to a system of sowing before the tropical wet season, then relying on fruit set during the dry season before harvest in winter [4]. This system and regional production collapsed due to heavy insect pressure and resistance by *Helicoverpa armigera* to insecticides. This program moved to northern NSW in 1972 and initially had strong emphasis on host plant resistance, with success in releasing okra leaf cultivars with tolerance to mites and *Helicoverpa* [5, 6] that set the foundations for future cultivar development by CSIRO. This program has grown rapidly since 1972 and has released many cultivars which now dominate the Australian cotton production system [7].

2. Challenges faced in breeding of Australian cultivars

Breeding challenges have largely been dictated by unique biotic and abiotic limitations on production and the need for Australia to remain competitive with other producing countries

that have lower labour costs. There has therefore always been a high focus on improving yield potential as a first priority and incremental increases in fibre quality to meet changing spinning requirements that might restrict the marketability of Australian cotton. Overlaid on these efforts has been targeted breeding for yield protection from losses caused from individual biotic and abiotic factors. Some of the early emphasis on breeding for host-plant resistance to insects has changed with the advent of genetically modified traits, but this has just moved the priorities to other challenges. A key asset to the breeding has been access to a diverse collection of germplasm from around the world, building on a quite diverse suite of base cultivars developed in the early 1970s and 1980s that have allowed CSIRO to tackle challenging insect, disease and weed issues.

2.1. Insects

A number of important pests of cotton are found in Australia that have necessitated the establishment of intensive pest management programs [8]. Cotton breeders have initiated programs to address host plant resistance and eventually released local cultivars containing *Bt* traits to confer resistance to the most important *Lepidopteron* pests.

The key pests of cotton in Australia are the larvae of two lepidopteron species, *Helicoverpa armigera* and *H. punctigera*, together with cotton aphid (*Aphis gossypii*), green mirid (*Creontiades dilutus*), spider mites (*Tetranychus urticae*) and silverleaf whitefly (*Bemisia tabaci* B-Biotype). From the 1960's to early 2000's, pest management was highly reliant on use of insecticides, mostly broad spectrum organophosphates, carbamates and pyrethroids as well as endosulfan. Crops were sprayed around 12-16 times per season [9]. This has brought with it predictable problems of pesticide resistance, destruction of natural enemy populations resulting in pest resurgence and outbreaks of secondary pests such as aphids [10] and spider mites [11], health concerns and off-farm movement into sensitive riverine environments [9].

Deployment of Ingard® (Monsanto's Bollgard) cotton in 1996 reduced insecticide use on those crops by about 50%, but efficacy was limited due to declining expression of *Cry1Ac* through the growing season. Ingard® cotton was always seen as an interim technology and during this period its area was capped at 30% to reduce the risk that *Helicoverpa spp*. would develop resistance to this critical insecticidal protein. This regulation limited the influence of the technology on the industry. The strong reliance on insecticides continued on the remaining 70% of conventional cotton and led to ongoing selection for resistance to insecticides in *H. armigera*, secondary pest problems and selection of pesticide resistance in these secondary pests. For instance, by the early 2000's spider mites were resistant to organophosphates, the pyrethroid bifenthrin and chlorfenapyr, and cotton aphids were resistant to organophosphates, the carbamate pirimicarb and pyrethroids [12].

The deployment of Bollgard II® in 2003/04, for improved resistance management and removal of the 30% area cap, resulted in a dramatic uptake of this technology. It also led to a massive decline in pesticide use (by about 85% [13]), especially relative to earlier years but also in comparison with contemporary conventional cotton that had also significantly reduced its reliance on pesticides. This technology essentially saved the cotton industry from insecticide

resistance in *H. armigera*, although a pre-emptive *Bt* resistance management strategy was initiated [12].

Over the past 10 years, silverleaf whitefly has occasionally reached pest status in some regions of the Australian cotton industry. It was first reported in cotton regions in 1994 [14], and the first major outbreak occurred in the Emerald production region in 2001/02. There are limited effective chemical control options for whitefly and they are also very expensive, so there is large emphasis on sampling protocols, thresholds and cultural control of this pest [15].

A range of other minor pests have occurred since 2000, many of which are related to the introduction and use of Bollgard II®. Reduced spraying for *Helicoverpa* has allowed survival of pale cotton stainers (*Dysdercus sidae*), jassids (*Austroasca viridigrisea*) and thrips (*Frankliniella schultzei* and *F. occidentalis*) late in the season and in some years these species have all caused damage and sporadically required control [12].

2.2. Weeds

In Australia, weeds are generally less of a problem in cotton than insects such as *Helicoverpa*, but many areas do have a reasonably high incidence of problematic weeds particularly on former grazing and/or flood-prone land [16]. Some of the problem weeds are *Ipomea spp., Cyperus rotundas, Polymeria spp., Conyze bonariensis,* and *Datura spp.* Weed control systems have changed in the last decade: production has changed from a cultivation-based system with residual herbicides and hand hoeing, to a system of minimal cultivation, with the use of herbicide tolerant cultivars (glyphosate and glufosinate) and few if any residual herbicides [17]. This has caused weed composition to shift to glyphosate tolerant species (particularly *Conyze sp.*) and has the real risk of developing glyphosate resistant weed populations.

2.3. Diseases

Many diseases of cotton are present in Australia and have necessitated the establishment of breeding projects to confer host resistance.

Bacterial blight of cotton caused by *Xanthomonas axonopodis* pv. *malvacearum* was historically the most important and widespread disease in Australian cotton. This disease was present in all crops and averaged around 20% of bolls affected though the mid-to late-1980s. The use of standard differential cultivars showed that Race 18 of the pathogen was predominant [18] and that infected seed was the major factor in the epidemiology of the disease. The disease was partially controlled through the production of blight-free planting seed. The CSIRO breeding program has released cultivars with immunity to blight and this disease is no longer found in Australian production systems due to 100% uptake of blight-resistant cultivars.

Verticillium wilt caused by *Verticillium dahliae* has been an important disease in Australia for many years, especially in the older growing regions of the Namoi, Gwydir and McIntyre. It is assumed that this pathogen is endemic to Australia. To date, all the Australian Verticillium isolates that have been tested are categorised as 'non-defoliating' strains [19]. Despite not having the more severe 'defoliating' strains that occur in other countries, yield reductions of

up to 20% have been recorded [20]. The majority of cultivars now grown commercially have relatively strong resistance to Verticillium wilt, but in seasonal conditions that favour the disease (temperature between 22-27°C [21]) and high number of rain days during the boll filling period [Allen S.J, pers comm.], significant productions losses still occur in some regions so improving resistance remains a breeding focus at a somewhat lower priority than other diseases.

Fusarium wilt caused by *Fusarium oxysporum* f sp. *vasinfectum*, was first recognised in Australia in 1993 [22] and is characterised by causing plant mortality at any time throughout the season, from seedling emergence through to harvest. By the end of 1999, the disease was present in six of the ten cotton production areas in Eastern Australia, but only widespread in two of those areas [23]. Subsequently, it has been shown that this is a unique strain of Fusarium and is substantially different to strains that are pathogenic to cotton in other countries [24]. This new Australian strain is identified as race 6 and is associated with alkaline clays, absence of nematodes [25], seedling death and an optimum temperature of 18 to 23°C [26]. Fusarium in the USA are races 1 and 2 and are associated with acid sands and nematodes [25], symptoms occur in mid-season with no plant mortality and an optimum temperature of 30 to 32°C [27]. It is assumed that the Australian pathogen evolved locally on weeds within our production system [28]. The extreme virulence and persistence of this pathogen and its ability to be readily transported in soil, water or trash raised concern within the industry when it was first discovered [29], particularly because most commercially grown cultivars in 1994 were highly susceptible. Production losses of virtually 100% have been reported for some fields. Substantial progress has been made in developing more resistant cultivars. However, for fields that have a high level of inoculum, the most resistant cultivar may still only have 10% survival in seasons that favour the disease so it remains a significant breeding challenge.

Cotton bunchy top disease (CBT) was first recorded in the late 1990s when a severe outbreak occurred in several Australian growing regions. Economic losses in that season alone were estimated at AUD$70 million [30], with nearly all Australian cotton cultivars being susceptible to the disease. The cotton aphid (*Aphis gossypii*) was identified as the vector and since then, there have only been sporadic outbreaks as the vector has generally been well controlled but its incidence is now very widespread across the industry. The causal agent has subsequently been identified as a Polerovirus and has some similarities to cotton blue disease found in Africa, Asia and the Americas [30]. Currently no commercial cultivars in Australia have resistance to this disease, although a readily accessible source of resistance has been identified in Australian material and cultivars are due for release around 2017.

Black root rot caused by *Thielaviopsis basicola* was first reported on cotton in Australia in 1990 [31]. It is now present in almost all cotton fields in NSW, as well as many in Qld. Diseased plants show stunted growth early in the season when the cooler temperatures favour the pathogen. While the pathogen does not kill plants, it can cause significant yield loss in cooler and shorter season regions due to the delay in crop maturation. The wide host range of the pathogen and persistence of the chlamydospores makes effective control through crop rotation very difficult. No *G. hirsutum* or *G. barbadense* cotton cultivars have been shown to have any significant resistance to this pathogen, so breeders must look to primary, secondary or tertiary germplasm or to GM approaches for control of this disease.

Alternaria leaf spot caused by *Alternaria macrospora* has caused severe defoliation of cotton grown on less fertile soils when the crop is exposed to extended periods of wet weather. These conditions occasionally occur in some Queensland growing regions. All upland cultivars currently grown have some degree of resistance, so the disease is considered of minor importance [20].

Reniform nematodes, *Rotylenchulus reniformis*, associated with stunting of cotton plants have only recently been identified in Australia [32] and are yet to be considered of economic importance. There was an isolated detection in 2003 in central Qld, but a more widespread identification was made in 2012, so it may become more important with time and pre-emptive breeding for resistance may be desirable.

2.4. Other industry challenges

A range of other industry challenges are also being, or need to be, addressed through breeding. There is a constant awareness of the dangers of having a narrow genetic base from which cultivars are developed. However, it is also acknowledged that broadening the genetic base through the introduction of exotic germplasm can often have a detrimental effect on yield [33]. In Australia, there is clear evidence of continuing yield improvement through selection within a 'narrowing' germplasm pool. In a high input, high yield industry such as Australia, there is a requirement for continued yield improvement to counter the ever increasing cost of production. While some of these improvements can come from improved management, historical data indicates that new cultivars are expected to provide at least 50% of the increase [2].

Many of the new challenges are related to the introduction of new technology or new environmental standards. The introduction of GM insect resistance traits into Australia in the mid 1990s simplified many aspects of pest management. However, it was soon discovered that the combination of high fruit retention of insect-protected crops together with early-maturity genotypes resulted in crops that had a fruiting cycle shorter than the available growing season and thus were not capitalising on the inherent yield potential of the cultivars. This had a detrimental effect on yield, so longer season, more indeterminate germplasm was subsequently used to deploy the GM traits. This remains a challenge as the industry expands into the southern (shorter season) regions of Australia where yield expectations of growers are similar to the fuller season traditional regions and may require a different approach to breeding.

Most Australian cotton is exported to Asia and spun on ring-spinning frames. Although this spinning technology has been used for around 100 years, modern machines continue to run faster and spinning companies demand greater throughput. This places greater stresses on the fibre, so there is a continual demand for higher quality fibre. There are well documented negative associations between yield and quality [34] which makes the simultaneous improvement of both yield and fibre quality difficult. In addition, breeders are also expected to address changing future fibre quality requirements due to changes in spinning technology or the yarn that is required by the textile industry.

At a time of increased awareness of the potential for climate change, greater scrutiny is also being placed on climate-related aspects of agricultural production [35]. Australia already has

a limited water supply, so increases in water-use efficiency (WUE) are being sought through management and breeding. In addition, the energy consumed and the emissions produced by a cotton crop are also being closely examined. Due to the heavy influence that yield has in the WUE calculation, this must continue to be the major focus of breeding programs. The efficient use of other resources, such as nitrogen fertiliser, will significantly influence the emissions from a crop [36, 37], so breeding programs also have responsibilities to continue research in this area.

3. Content and description of Australian cotton germplasm collections

Initially, Australian plant breeders acquired and maintained their own germplasm collections. However, in the early 1980s the Commonwealth and State governments established a series of genetic resource centres to conserve national germplasm collections. The Australian Tropical Crops and Forages Genetic Resource Centre (ATCF) Biloela, Qld is the centre that included cotton. Currently, the only other two cotton collections in Australia reside at CSIRO Plant Industry in Narrabri, NSW and Canberra, ACT.

Unfortunately, the ATCF cotton collection is, at present, only funded to maintain germplasm under long-term storage, so the collection is not actively maintained or expanded. However, germplasm can still be requested from this collection. The CSIRO collection at Narrabri is maintained and funded by the dedicated breeding program. The CSIRO collection in Canberra is largely dedicated to the long-term storage of indigenous Australian species that have been collected through germplasm explorations [38] over the last few decades. Table 1 describes the number of cultivated tetraploid accessions amalgamated across the ATCF and CSIRO Narrabri collections. This includes both cultivars as well as locally developed and introduced germplasm. Table 2 describes the wild accessions held in all three collections.

Species	Description	Number of accessions
G. hirsutum	Germplasm	952
	Cultivars	611
	Landraces	10
	Total	*1573*
G. barbadense	Germplasm	51
	Cultivars	45
	Landraces	3
	Total	*99*
Other hybrids/genetic stocks		12

Table 1. Number of cultivated tetraploid accessions in Australian collections.

Species	Number of accessions	Genome*
Primary gene pool		
G. tomentosum	6	AD3
G. mustelinum	1	AD4
G.darwinii	1	AD5
Secondary gene pool		
G. herbaceum	39	A1
G. arboreum	211	A2
G. anomalum	4	B1
G. logicalyx	1	F1
G. thurberi	5	D1
G. trilobum	1	D8
G. davidsonii	2	D3-d
G. klotzschianum	1	D3-k
G. armourianum	1	D2-1
G. harknessii	1	D2-2
G. aridum	1	D4
G. raimondii	1	D5
Tertiary gene pool		
G.sturtianum	64	C1
G.robinsonii	3	C2
G. stocksii	3	E1
G. somalense	2	E2
G.australe	159	G
G.nelsonii	39	G
G.bickii	28	G1
G.costulatum	1	K
G.cunninghamii	2	K
G.enthyle	1	K
G.exiguum	3	K
G.nobile	1	K
G.pilosum	1	K
G.populifolium	1	K
G.pulchellum	1	K
G.rotundifolium	3	K
unclassified	8	

* As per [39]

Table 2. Number of wild primary, secondary and tertiary tetraploid and diploid accessions in Australian collections.

3.1. Australian native *Gossypium*

Australia is home to 17 native *Gossypium* species that are found exclusively in Australia and are distantly related to cultivated cotton. These contain the C, G and K genome species. The distribution and characterisation of these diploid species has been reasonably well documented [40-43]. The potential for transferring traits from these species, together with techniques for the production of fertile hybrids has also been reported [44, 45]. There is interest in a number of traits from these species, however, research has largely focused on disease resistance [46, 47], glandless, and seed-glanded plants [45, 48]. Despite this research, there are no documented cases of successfully transferring economically important traits from the wild Australian *Gossypium* species to cultivated cotton. There still remains interest in this area and improvements in molecular techniques may facilitate the successful introgression of traits.

3.2. Plant quarantine requirements

Australia has strict plant quarantine regulations which aim to safeguard Australia's plant health status to maintain overseas markets and protect the economy from the impact of exotic pests and diseases. This is administered by the Commonwealth (Federal) Department of Agriculture, Fisheries and Forestry (DAFF) Biosecurity based on policy determined in consultation with industry stakeholders (http://www.daff.gov.au/bsg) [verified December 2013]. For cotton, these requirements are quite onerous and costly, which restricts capacity for importing cotton seed into Australia.

3.3. Focus of the collections

Both the ATCF and the CSIRO Canberra collections have actively collected wild *Gossypium* accessions. As can be seen in Table 2, this has largely focused on *G. sturtianum, G. australe, G. nelsonii* and *G. bickii.* However, the major focus of the CSIRO Narrabri and ATCF collections have been the collecting of overseas cultivars, predominantly *G. hirsutum*, for use in the Australian cotton breeding programs. Over time, these programs have had a range of objectives as mentioned previously which is reflected in the material accumulated. Major areas of interest have been general genetic diversity, targeted host-plant resistance to insects and diseases, crop maturity, fibre quality and yield.

3.4. Key contributors to collections

Over time, germplasm representing most cotton producing countries has been sourced. This has either come directly from breeders (through exchange), acquisition of commercially released cultivars, or requests to centralised cotton collections. Initially, germplasm was sourced in larger numbers to rapidly develop the collections, however, as time has gone on, the requests for germplasm have been much more focused, as for example when Fusarium wilt was discovered in the early 1990s in Australia. Although this is a locally evolved strain that is unique to Australia, significant efforts went into searching for germplasm that had improved levels of resistance. Extensive literature searches were undertaken to identify potential germplasm from other countries to evaluate for resistance. Once a genotype had been

identified, various databases were searched and institutions contacted in an attempt to find a source for the material. Although germplasm was sourced from a number of countries, the USDA collection in TX proved to be an extremely valuable resource for resistance to Australian isolates of Fusarium. This example highlights the need to be aware of what is available, to seek out that germplasm and request it under as favourable terms as possible.

3.5. Current status of collections

Although the ATCF collection is acknowledged as a valuable national resource, funding of this collection remains uncertain. On the other hand, the CSIRO collection at Narrabri is a live collection associated with an active breeding program. Although not an extensive collection by world standards, containing around 1000 introduced cultivars and germplasm, it is nonetheless a critical component of the CSIRO breeding program. The collection is continually being added to when new material is identified which may contain diversity for traits of interest. Many of these relate to insect and disease resistance, but yield, fibre quality and maturity are also targeted.

3.6. Storage and renewal of germplasm collections

The ATCF collection is maintained under long-term conditions in which germplasm is preserved as fuzzy seed that is dried to 6% moisture content at 15°C. Once the seed is dry, it is sealed in laminated aluminium bags and stored at -20°C. Under these conditions, it is expected that viability will be maintained for at least 50 years. Previously, when funding was available, all accessions were routinely monitored for germination every 10 years and regenerated if germination fell below 60% of the starting germination. Seed was either regenerated in the field or greenhouse with controlled pollination. Germplasm was extensively characterised in the past, but this is no longer undertaken. The collection maintains an in-house seed management database that manages germplasm passport, collection, characterization, inventory, seed health, regeneration and distribution data. The passport and characterisation data are freely available via the Australian Plant Genetic Resources Information System Web site (http://www2.dpi.qld.gov.au/extra/asp/auspgris) [verified December 2013]. A national review of Plant Genetic Resource Centres in Australia has determined that this website will be replaced with GRIN-Global. However, a deadline for this transition has not been finalised.

The CSIRO Narrabri collection is maintained in short-term storage as delinted seed at 7°C and 50% relative humidity. Around 1 kg of each accession are stored in cotton bags. A 10g backup collection is also maintained in sealed laminated aluminium bags at 4°C in an additional facility on the same site. Each accession is regenerated in the field every 10 years, or when stocks have been depleted to less than 200g. During regeneration, accessions are inspected and compared to known visual characteristics and off types are removed. Accessions that are photoperiod sensitive are grown in the greenhouse during the winter when the short days promote flowering. In addition, those accessions in which purity is considered vital to maintain are also grown in the greenhouse. The CSIRO breeding program maintains an internal database where accession data is recorded. This database is not publicly available.

The CSIRO Canberra collection is in long-term storage at -20°C. This collection provides a resource, but is not currently funded or maintained. Many of its accessions are also available from major US germplasm collections, as US researchers were often involved in the initial exploration. As the majority of these lines have been collected during genetic explorations, an internal database contains details of the collector, location and date. This database is not publicly available. There has been a decline in funding in CSIRO over the last few decades for extending germplasm collections for native *Gossypium* and other crop relatives like *Glycine*, as more and more biodiversity research shifts to consolidation of national collections, electronic archiving of data and making them more broadly accessible on-line.

There are increasing challenges associated with maintenance and regeneration of these collections. Ensuring genetic purity of accessions has always been difficult when regeneration is conducted in the field. However, the GM era has added another level of complexity to the problem. Even though Australia does not have large numbers of insects that will cause cross pollination in cotton (only the European honey bee), measurable out-crossing can occur in some seasons and some locations [49, 50]. Care must be taken to locate regeneration blocks away from known sources of bee activity as well as isolated from commercial GM cotton crops. CSIRO also has substantial activities in areas of biotech research in addition to the traits sourced from third parties and some of these activities are discussed in a later section. This research is geographically isolated (in Canberra) from the breeding program and only traits that have reached the advanced field evaluation stage are grown at Narrabri. As CSIRO is developing cultivars for release commercially, stewardship of germplasm and GM traits is extremely important and internal quality assurance protocols are rigorously adhered to. These protocols dictate that all material needs to be tested via protein or DNA analysis prior to handover to commercial partners for seed increase and sale.

3.7. Germplasm passport data

The current status of the accession descriptions varies depending on when and who imported them, as well as the donor. As many characteristics are either invisible or not able to be determined except in specific environments (e.g., disease resistance), the data provided by the donor is often critical. However, in many cases, particularly for older material, this information is simply not available for many accessions that may have been obtained via an intermediary collection or supplier. When it was actively funded, the ATCF collection generated substantial passport data on all lines including morphological characteristics, boll size, lint percent and fibre quality parameters. This is invaluable information for a breeder searching for specific characteristics. Descriptions in the CSIRO collection are more *ad-hoc*. This is largely a function of the collection not being 'public', and simply a resource for the breeding program that operates it. The data that does exist largely relate to the traits that were initially identified as being of interest *i.e.*, HPR to insects and disease or fibre quality.

3.8. Sharing and exchange

The ATCF collection is based on a national collection model with small quantities (20-30 seeds) of germplasm freely available to researchers worldwide. The CSIRO collection is not publicly

available, but historically, genuine exchange with other breeding programs has been practised. In recent years however, commercial agreements and patents have made this exchange difficult. This has restricted the amount of new material added to the collection. However, it is imperative that exchange of germplasm continues. Many of the patent and commercial 'in confidence' issues can be addressed with specific (though often restrictive) material transfer agreements. This can allow new germplasm to be evaluated prior to negotiating further freedoms to utilise germplasm in a breeding program.

4. Other cotton germplasm resources beyond seed collections

4.1. Mapping populations

Improving disease resistance against indigenous Australian strains of Fusarium has been a high priority in the Australian breeding program and a number of new sources of resistance have been identified through the screening of introductions (mostly) from the US, China and India where Fusarium is also endemic. To investigate the genetic basis of the Fusarium wilt resistance in the Indian *G. hirsutum* cultivar MCU-5, a bi-parental cross was made between MCU-5 and Siokra 1-4; a *Fusarium* susceptible Australian okra leaf *G. hirsutum* cultivar [51]. An F_3 population consisting of 244 lines was developed from this cross, and from single seed decent from each F_3 line, 244 F_4 lines subsequently produced. The F_3 and F_4 populations were assayed for Fusarium wilt resistance using a glasshouse bioassay [52] and genotyped with 151 markers (95 SSR and 56 AFLP). QTL analysis revealed the presence of multiple regions that were associated with resistance that provide targets for introgression into elite cultivars to improve Fusarium wilt resistance [51]. Subsequently, it was found that MCU-5 which was derived from a multi-line cross between Indian Cambodia-type cultivars (MCU-1 and MCU-2) and cultivars from East Africa, the West Indies and the US, including some contribution from *G. barbadense*, also possesses significant resistance to non-defoliating strains of *Verticillium dahliae,* as well as possessing significantly longer fibre than Siokra 1-4.

In an international collaboration with CIRAD (France), CSIRO obtained a 140 RIL inter-specific population ranging from the F_6 to F_9 stages of selfing through single seed descent. The two parents; Guazuncho 2 (*G. hirsutum*), and VH8-4602 (*G. barbadense*), were chosen for their agronomic performance (Guazuncho 2) and superior fibre quality parameters (VH8-4602) [53]. This population was grown in several countries including Australia in the glasshouse and field and analysed for many different traits (including fibre, leaf shape hairiness, boll size and number and earliness) and genotypes consisting of 1,745 loci derived from 597 SSR and 763 ALFP markers [54-56]. Not all of the lines performed well under Australian conditions with only about 55 flowering within an acceptable timeframe for harvesting and thus seed stocks are severely restricted for most of the RILs. The QTLs associated with fibre traits identified in Australia may help generate markers for selection for these traits in breeding populations. The population also represents a resource for linking markers with mite resistance, as the parent lines show differentiation in resistance.

4.2. Near isogenic lines of host plant resistance traits

Cotton mutants can be of interest to cotton breeders for their agronomic or host plant resistance possibilities. Near Isogenic lines (NIL) are plants that mostly genetically identical except in DNA regions associated with a specific trait or gene mutation. They are usually created by the repeated backcrossing of a mutant plant to a recurring parent, and are useful for quantifying the effect of the mutation or phenotype on agronomic performance, by reducing genetic background differences between the mutant and the normal plant type. Although very important for defining the agronomic value of an altered phenotype, NILs require a large investment of time to produce. Thomson [57] developed NILs for glabrousness (T_2^{arm}), frego bract (fg), okra leaf (L_2^0), and nectariless (ne_1,ne_2). These mutations are known to affect resistance to specific insects and diseases, and in the case of glabrous also reduces lint trash [57]. All 16 NIL combinations of the four mutants were developed in the Deltapine 61 background. The 16 NIL were derived from an initial crossing of a Deltapine-related experimental line homozygous for all four mutant characters and Deltapine 61. Four backcrosses were then made with Deltapine 61. A final backcross was made and from a large F_2 population homozygous genotypes were selected and maintained. Subsequent hybridizations were done using lines from this original population as parents and a total of nine cultivars were released up until 1999 including the original Siokra and Sicot cultivars.

4.3. Mutant populations

Mutation and mutation breeding is a tool for producing novel variation that potentially cannot be found in existing cotton species, for genetic improvement and in aiding the study of gene function. Mutations are nucleotide base changes within the genome of an organism that are not brought on by normal recombination and segregation, and occur naturally at a low frequency. Many naturally derived mutant cotton plants have been isolated that contain no fibre, which has aided our molecular understanding of fibre formation [58]. The rate of mutation can be greatly increased or induced in plants using chemical mutagens, ionizing radiation or transposable elements. Induced mutants have been generated and used to create valuable traits in many major crops, but have only occasionally been used in improving cotton. Mutagenesis of cotton has resulted in 'naked and tufted' seeds, herbicide resistance and plants with longer fibre [59-62] that have direct application within the cotton industry. Other mutants, such as those possessing inferior fibre traits, provide powerful tools for understanding what genes are associated with fibre formation [58].

CSIRO is investigating the general usefulness of mutagenised cotton populations for conventional plant breeding. Populations of *G. hirsutum*, *G. barbadense* and *G. arboreum* have been mutagenised with the chemical mutagens sodium azide and ethyl methanesulfonate, as well as heavy-ion mutagenesis using the Riken Ring Cyclotron at the RI Beam facility in Japan in collaboration with Dr Tomoko Abe. The mutagenised populations are currently being screened for a number of traits including: herbicide resistance, fibre traits, boll size, plant architecture and flowering.

http://dx.doi.org/10.5772/58414

4.4. Transgenics

A transgenic cotton plant contains gene/s that have been artificially inserted. The inserted gene (known as a transgene) may come from another cotton plant, or from a completely different species, such as *Bt* cotton; which possesses the *Bt* toxin from the bacterium *Bacillus thuringiensis*. Transgenic technology enables plant breeders to bring together in one plant traits from potentially any organism, not just from within Upland cotton or sexually compatible species. Potentially this technology could be a quicker and cleaner method of bringing in traits from poor agronomic sources compared to backcrossing, but regulatory costs makes this currently unfeasible. Transgenic technology also provides the means for studying specific genes in cotton and enables the definitive assignment of genes to specific functions through either over-expression or silencing of individual genes or gene families.

It is difficult and time consuming to generate transgenic cotton, as it requires *Agrobacterium tumefaciens* mediated gene insertion via callus, generated via tissue culture. It was found in the 1980s that Coker cultivars were generally superior for gene transfer using tissue culture, although a small number of Australian cultivars could be transformed with relatively low efficiency [63]. Coker 315 is the cultivar used for all transformation events in Australia as it was thought to possess better agronomic traits under Australian conditions compared to other transformable Coker cultivars, such as Coker 312 that have been used by international biotech companies. Currently, all transgenic traits present in Australian commercial cultivars were obtained internationally under License from either Monsanto or Bayer CropScience. The focus of Australian transgenic R&D work has been on developing potential traits of more specific relevance to the Australian Industry and on basic research to understand gene function, particularly the functions of genes involved in fibre initiation and formation, plant defence and seed oil formation. Such work has demonstrated the importance of GhMyb 25 [64], GhMyb25-like [65] and GhHD-1 [66] for fibre initiation and for controlling fibre initial numbers. As part of CSIRO research, transgenic cotton plants that are altered in the expression of GhMYB109 were imported under a Material Transfer Agreement for research use [67]. CSIRO transgenic research has demonstrated through altering the expression of a cotton sucrose synthase [68] or overexpressing a potato sucrose synthase [69] the essential role of the different sucrose synthases in cotton fibre and seed formation. Expression of the *Talaromyces flavus* glucose oxidase gene in cotton demonstrated increased resistance against Verticillium wilt [70], and plants with suppressed Cadinene synthase gene expression revealed its role in bacterial blight infection [71]. Although cotton is principally a crop grown for fibre, its seeds are also a valuable source of oil. Transgenic silencing of key fatty acid desaturase genes altered the cottonseed oil composition to improve its nutritional profile, making it more competitive with other oilseed crops has also been achieved, although commercialisation of this trait is still under discussion [72]. Currently none of these CSIRO transgenic derived traits have been incorporated into commercial cultivars.

Successful expression of transgenes in cotton also requires the ability to precisely express genes at high levels. Experiments with transgenic plants have demonstrated the value of the soybean lectin gene promoter to drive transgenes in the embryo of cotton seeds [73], and subterranean clover stunt virus promoters and terminators [74] for general and high level expression in

cotton. The CSIRO derived sub-clover duplicated stunt7 viral promoter was deregulated in 2008 as a commercial event in cotton (T304-40) driving the *Cry1Ab* insect resistance gene and forms part of Bayer CropScience's TwinLink product. A cotton rubisco small subunit promoter has also been shown to be expressed at high levels in green photosynthetic tissues throughout the development of cotton in the field and hence a useful promoter for expressing transgenes in leaves [75].

5. Molecular markers for germplasm characterisation

Molecular markers are DNA tags or sequence differences (polymorphisms) that provide the ability to quickly track the presence of specific DNA regions associated with a trait, through a segregating population. They are widely used for basic research in plant biology and genetics and are now essential to the CSIRO cotton breeding program for selecting homozygous transgenic lines on a large scale, monitoring the purity of transgenic breeding lines, and stacking traits by marker assisted selection. Upland cotton has limited intra-specific genetic polymorphism compared to other crops as revealed by a number of molecular marker and genetic diversity analyses [76-78], due to the relatively recent polyploidisation (estimated 1-2 mya) [79, 80] that created the species as well as domestication and selection. This has hindered application of molecular markers to breeding, as to obtain large numbers of polymorphic markers for detailed mapping required crossing *G. hirsutum* to *G. barbadense* (inter-specific), which are largely unsuitable for cultivar improvement in breeding programs.

Two major technological breakthroughs have occurred in the marker field since ~2010 that have removed many limitations associated with marker technology and its application in cotton.

1. Next-generation sequencing (NGS) can sequence DNA millions of more times than previous methods [81]. This has opened up the possibility to discover a larger number of single nucleotide polymorphisms (SNP, single DNA base changes) in cotton that are the most common form of difference between individuals or cultivars, and should be sufficiently abundant to discriminate between any two *G. hirsutum* lines. NGS has also enabled the completion of the genome sequence of the diploid cotton *Gossypium raimondii* [82] that is related to the D-genome present in Upland cotton. The genome sequence of a diploid A-genome containing cotton species is likely to be publicly available in the near future, and will make it easy to compare and align any short sequence reads from NGS to find large numbers of widely distributed nucleotide differences between cultivars that can be used for mapping and breeding.

2. New high-throughput genotyping (HTG) platforms based on SNPs have been developed that can accurately call millions of SNPs in large populations [83, 84]. These technologies enable complete genome coverage of markers between *G. hirsutum* cultivars in a fraction of the time required by conventional marker approaches.

5.1. SNP discovery in germplasm important for Australia

To ensure that SNPs are informative to the breeding populations being developed in Australia, it is essential to find SNPs among cotton cultivars that constitute the major germplasm sources of the elite cultivars we are developing. The most straightforward method to identifying SNPs, in the absence of the Upland cotton genome sequence, is to sequence expressed gene transcripts (RNA-seq) by isolating mRNA and converting it into cDNA for sequencing. This method has been used successfully for many other plants such as maize and wheat [85]. RNA-seq targets SNP discovery to genes that are actively transcribed and therefore more likely to be associated with conferring trait differences. The disadvantage is that cDNA is likely to have lower SNP frequencies than non-expressed regions as they are constrained by the genes function. CSIRO RNA-seq data was generated on a set of 18 cultivars that represented significant genetic variation present within current Australian commercial cultivars were selected; containing old Australian and US cultivars, as well as cultivars from China and India. Over 50 million reads (of ~90 bp) for each Upland cotton sample was obtained. We found the key to identifications of varietal SNPs confidently was when a sub-genome-specific SNP was also found in close proximity to the varietal SNP (Figure 2). This enabled representative sequences from both genomes in each cultivar to be identified and compared. From the 18 cultivars ~38,000 varietal SNPs were identified. A selected subset of >1,500 of these putative SNPs were analysed using a combination of SNP platforms (GoldenGate and Sequenom) and it was found that these SNP could be validated at a rate >90% [Zhu, Q-H, pers comm.].

```
Cultivar-1A  ……AGCGTAGTCAGATTAAGTGGAATCCCTGATG……
Cultivar-1D  ……AGCCTAGTCAGATTGAGTGGAATCCCTGATG……
Cultivar-2A  ……AGCGTAGTCAGATTGAGTGGAATCCCTGATG……
Cultivar-2D  ……AGCCTAGTCAGATTGAGTGGAATCCCTGATG……
```

Figure 2. Stretch of DNA sequence from two cultivars showing the sequences for both the A and D genomes. SNPs with the best validation rates are where a varietal SNP (in red) is in close proximity to a sub-genome-specific SNP (in blue) that allows determination of which genome the specific short reads sequences are derived from.

Although RNASeq data has allowed us to progress towards being able to effectively use SNPs for genotyping in breeding projects, the protein coding regions of genomes have been found to have a significantly lower level of DNA polymorphism than non-coding regions, so polymorphisms within genes between closely related cultivars are going to be less frequent and hence less useful. With the availability of the assembled *G. raimondii* genome and possibly of the *G. arboreum* genome soon, to serve as a framework for short read sequence alignment, our SNP identification will in future be performed using genomic DNA sources.

5.2. International cotton SNP consortium

SNP Chips enable high throughput parallel analysis (millions at a time) whereas older markers like SSR markers tend to be performed only one or a few at a time. Therefore SNP Chips represent a significant improvement over older technologies for large-scale genotyping.

Recently an international cotton consortium was formed to create a 70,000 public Illumina Infinium SNP array for cotton. This array was made available for purchase in late 2013 and contains ~ 50,000 intra-specific *G. hirsutum* SNPs, ~16,000 inter-specific SNPs predominantly from *G. barbadense* but also *G. tomentosum* and *G. mustelinum*, and small numbers (~4,000) of SNPs from two diploids *G. longicalyx* and *G. armourianum*. The publicly available SNPs were provided by a number of international groups including; CSIRO, Texas A&M, University of California-Davis, Cotton Incorporated, Brigham Young University and United States Department of Agriculture-Agricultural Research Service, Centre de Coopération Internationale en Recherche Agronomique pour le Développement (CIRAD), Council of Scientific and Industrial Research-National Botanical Research Institute (CSIR-NBRI), and Dow AgroSciences. These arrays will provide unprecedented numbers of makers to be screened across cotton germplasm and will likely be the genotyping method of choice for a number of years.

5.3. Genotype-by-sequencing

SNP Chips represent a major advance in genotyping of cotton. However, SNP Chip platforms still possess limitations. Since cultivated cotton is an allotetraploid it is often difficult for these platforms to differentiate between the two cotton sub-genomes, and from our experience only ~34% of the polymorphic SNP markers act as co-dominant markers (can discriminate between both homozygous alleles and their heterozygote class, see Figure 3A) on a GoldenGate or Fluidigm platform, whereas the majority act as dominant markers (can only differentiate accurately between both homozygote classes but the heterozygote class cannot be differentiated from one of the homozygote classes, see Figure 3B). Dominant markers reduce the amount of information that can be obtained from an individual, but they are still very useful for mapping (eg., commonly used AFLP markers are all dominant markers). In addition SNP platforms can only interrogate the SNPs that have already been indentified and placed on the chip, and once manufactured, new SNP Chips are unlikely to be remade for several years, so any newly discovered SNPs will be unable to be included and assayed.

We are investigating genotyping using NGS alone, which is called Genotype-by-Sequencing (GBS). GBS uses an amplified subset of genome from individual lines or plants to identify base differences between them [86, 87]. Using bioinformatic analyses the SNPs are found on-the-go when comparing the two sets of sequences, so no prior information about the genotypes is required. The SNPs found are then analysed as separate markers. The advantage of GBS is that SNPs are found in the analysis of the DNA fragments compared, and so informative SNPs linked to a trait are more likely to be identified. Also SNPs from different genomes can be easily separated based on the presence/absence of specific sub-genome related SNPs (which are much more common than varietal SNPs) and so most markers can be selected to be co-dominant (i.e., heterozygous alleles can be scored). This makes this genotyping technology especially suitable for polyploids such as cotton, and is becoming the genotyping method of choice in, for example, wheat which has three similar genomes (A, B and D) [86].

http://dx.doi.org/10.5772/58414

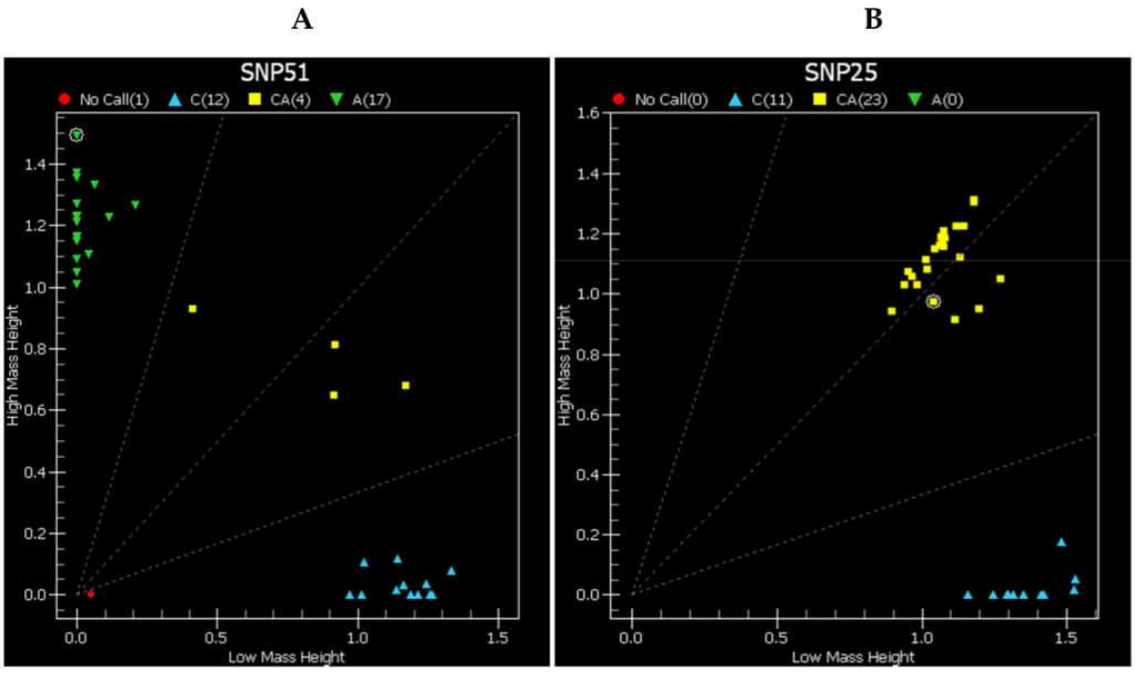

Figure 3. Typical *G. hirsutum* varietal SNP profiles using the Sequenom platform. Each spot represents a different cultivar being assessed for the presence of a specific SNP. A) Co-dominant example, a particular polymorphic nucleotide is scored either as homozygous A (Green) or C (Blue) or heterozygous for C and A (Yellow). B) Dominant example, a particular polymorphic nucleotide can only be scored either as homozygous A (Blue) or heterozygous CA (Yellow), homozygous C (blue) cannot be scored due to tetraploid nature [Zhu, Q-H, pers comm.].

5.4. Genetic diversity analysis

Possessing significant genetic diversity within a breeding program is extremely important for future crop improvement. Where detailed pedigree information is known, breeders can specifically select parents for crossing based on their trait package and degree of relatedness. However, it is often the case that detailed knowledge about the pedigrees of cultivars is lacking, especially for imported cultivars, and this may hamper their use in cotton improvement. Molecular markers provide an alternative means of determining levels of relatedness and ancestry between cultivars. Many studies have used markers to determine the diversity levels with cotton populations [88], however most have suffered from insufficient numbers of marker required to generate accurate estimates. The first major use of a large public cotton SNP chip and other high throughput genotyping technologies by CSIRO will be to enable more accurate diversity estimates within breeding populations and seed repositories. Breeders should be quickly able to determine diversity estimates and phylogenies between cultivars, as well as identify shared genetic regions between members of the same pedigree. This data will enable the use of more diverse germplasm which may contain unique traits to be effectively integrated into the breeding program.

5.5. Linking SNPs to traits

With improvements in SNP identification and the availability of large scale public cotton SNP chip, the availability of molecular makers will no longer be a limiting issue in cotton research and the main task will be linking markers to important traits that will aid in their introgression into elite varieties. Most of our initial focus will be to link SNPs to disease resistance traits, as increased disease resistance has been responsible for a significant amount of our variety based yield improvements. Common methods for linking markers to traits involve the creation of bi-parental crosses that segregate for the trait of interest, followed by selfing, genotyping and trait assessment in the F_2 or F_3 populations. Where possible immortal genetic populations, in which each line is inbred and can be propagated simply by self-pollination, such as recombinant inbred populations or inbred backcross (IBC) populations are used as phenotyping is much simpler (only looking at homozygote classes) and repeatable. These immortals lines have been very helpful when assaying for disease resistance (Verticillium wilt, Fusarium wilt and black root rot) as disease phenotypes are often environmentally affected and difficult to score. The new high-throughput marker technology should greatly accelerate these traditional genetic studies in cotton. However, such populations require significant amounts of time and labor invested in their creation. With the availability of high throughput SNP genotyping between *G. hirsutum* cultivars, other approaches consistent with conventional breeding strategies can now be routinely performed. One such method involves locating chromosomal segment substitutions derived by repeated backcrossing of the trait into an elite cultivar. Backcrossing is used by breeders to transfer a limited number of desirable traits from one parent (possibly poor agronomically) to an elite parent, and involves multiple rounds of crossing back to a recurring elite line; with each backcross reducing the amount of donor DNA present in the offspring. This methodology can be used to find markers linked to the trait(s) after as little as 3 cycles of backcrossing and then selfing, as individuals with the desired trait can now be compared using whole genome SNP analysis to reveal which genomic regions are associated with the donor parent. The markers in the donor regions can be confirmed as being linked to the trait in a further cycle of backcrossing, selfing and selection. Once the marker linkage has been verified, marker-assisted backcrossing can then eliminate the need for further trait phenotyping until the desired elite cultivar is obtained. We are using this system to find linked markers for disease and fibre traits and for introgression of traits from other cotton species.

The ability to scan the whole genome quickly and identify regions from the donor plant makes this strategy practical in cotton and is limited only by reliability and robustness of the trait and the time it takes to cycle through multiple generations. An added advantage of this method to find markers is that the germplasm produced at the end of the process is already partly introgressed into an elite cultivar and can be a parent for more advanced crosses with other elite material. A SNP Chip can also be used to identify backcrossed individuals with the lowest levels of donor DNA at each backcrossing step, reducing the number of backcrosses required to produce an elite cultivar with the trait of interest. This is especially useful to separate the trait of interest from regions nearby that may carry genes that result in poor yield (linkage drag), such as Okra leaf trait.

5.6. Genome wide association mapping

Large scale genotyping of cotton at an economical price enables genotyping large diverse collections of *G. hirsutum* lines. To link SNP markers to important traits in cotton the most manageable route is to perform genome wide association studies on elite cultivars and their pedigrees, avoiding the time and energy required in the creation of specialised genetic populations examining small numbers of traits. Association genetic analysis is a method for linking specific markers with phenotypes using established populations of individuals, and has been extensively used in human/animal genetic analyses where defined genetic crosses are often difficult to generate [89]. The extent of genome-wide allelic association (linkage disequilibrium: LD) is the key starting point for association mapping. The extent of LD has been quantified and association mapping has been successfully applied for many plant species (e.g., [90]) including cotton [91, 92]. The application of LD-based association mapping for cotton will facilitate comprehensive utilization of existing genetic diversity conserved within cotton germplasm cultivars.

CSIRO is interested in using association analysis on a large number of well defined cotton varieties to find markers linked to disease resistance, fibre quality and yield component traits. The bottleneck for this type of research is the difficulty in adequately phenotyping a large number of cultivars for agronomically important traits. To obtain an accurate measurement of the most important trait, yield, requires significant replication both within a field, across geographical regions and over multiple seasons. This restricts the numbers of varieties that can be tested lowering the power of the analyses. It is also possible to model genetic/environmental (GxE) interactions for specific traits using association analysis, but that requires a magnitude larger level of replication under different management conditions, as well as significant knowledge about the environmental conditions in which the traits were measured (and therefore is unlikely to be undertaken by our research teams in the near future).

5.7. Mining diversity in other cotton species

The cotton genus (*Gossypium*) contains ~50 species that are widely geographically distributed from arid to tropical regions and morphologically diverse ranging from herbaceous perennials to small trees. These species are usually divided into three gene pools with regard to their use for genetic improvement of Upland cotton though only a few sources have been extensively used. The primary gene pool consists of *G. barbadense*, *G. tomentosum*, *G. mustelinum* and *G. darwinii* which are tetraploid species with the same A and D genome complement that are sexually compatible with Upland cotton. The secondary gene pool is represented by diploids with A, D, B or F genomes that require synthetic tetraploid formation or a synthetic hexaploid bridging species for traits to be introgressed into Upland cotton [39, 44]. The tertiary gene pool consists of other diploid species with a completely different genome type such as C, E, G or K that show relatively poor or no recombination with the A or D-genome and thus traits from these species are likely to require isolation of the causal genes and transfer into cultivated Upland cotton by genetic modification.

Once a representative A-genome diploid sequence is completed, and its information added to the already published D genome, it will be possible to use this information as a scaffold to

determine the gene coding regions for different *Gossypium* species. The coding sequences from these genomes will identify new diversity within genes already associated with disease and stress that could be targets for introgression. Traditionally the work required to introgress traits from anything other than *G. hirsutum* or *G. barbadense* has been prohibitive. However the ability to identify important genes via re-sequencing of other *Gossypium* species means that markers for these traits can be found rapidly and used in marker assisted backcrossing strategies as mentioned above. Therefore, in future we will be able to access much more diverse germplasm for unique traits with potential targets for introgression including glandless seeds (in a glanded plant) and Fusarium wilt resistance that are present in indigenous Australian *Gossypium* species [47, 48].

5.8. Exploiting mutant collections of cotton

As previously mentioned, induced mutations can produce novel variation that can be directly used for crop improvement and aid in identifying which genes are important for specific traits. Now that a genome sequence is available for cotton, the major basic research challenge is to uncover the genes responsible for major agronomic traits from the more than 50,000 genes present in this species. The advent of NGS technology and the completed *G. raimondii* genome sequence has opened up practical ways of achieving this via a reverse genetics approach; where a mutation in a specific gene sequence is first identified in a plant by sequencing and the phenotype of this plant is then compared against the not mutated or wild-type plant to identify the physiological, morphological or agronomic consequences of the disruption of that gene and hence its possible function. Traditionally mutated plants with interesting phenotypes are selected from large populations of mutagenised plants, and researchers work back to find where the mutation that caused the novel trait is located within the genome (forward genetic approach). However the space/land required in order to select for interesting mutants is substantial for a large plant like cotton, and the tetraploid nature of Upland cotton, potentially makes finding valuable mutants more difficult (as for some traits both the A and D genome genes would need to be simultaneously mutated to result in a phenotype). The reverse genetics approach, is to first obtain a large collection of plants carrying chemically or radiation induced mutations spread throughout the genome that are identified through NGS. Since NGS can rapidly re-sequence a genome as long as an existing reference genome is known, sequencing the randomly mutated plants will identify single base changes in gene sequences caused by chemical mutagenesis similar to how SNPs are already identified between cultivars. Using specialised DNA capture technologies that only isolate DNA from expressed regions of the genome to reduce the sequence complexity [93, 94], this procedure would be cost effective. A large library of plants harbouring these known mutations, would be made by subjecting cotton seeds to induced mutagenesis and plants (M1 generation) grown and allowed to self seed. The M2 seeds would then be planted, leaf material isolated and the gene sequences of these plants sequenced and the multiple mutations in each plant identified. By collecting seeds from the M2 plants, a database of the sequence changes would be identified and associated with specific seed lots and mutations causing loss of gene function of any selected gene could be identified. The M3 seeds from that plant would be planted and using markers specific for the mutation (a type of SNP marker), plants selected that were homozygous for the mutation [95]. The

phenotype of these plants would then be compared to homozygous wild-type plants at this location from the same seed lot.

The availability of mutant lines in most or all of the cotton genes would provide the cotton community with the means of testing gene function in a large number of genes at an unprecedented rate. Lines with mutations in a particular gene would need to be found in both subgenomes, then crossed and progeny selected that were mutated for both copies of the genes in order to categorically define its function. Once a gene has been confirmed as being associated with an important trait, the sequence of this gene could be sequenced from many different cotton cultivars and related species to identify variants that may possess superior qualities for targeted introgression. Currently, large mutagenic populations of *G. arboreum* and *G. hirsutum* are being established at CSIRO in order to locate genes involved in fibre formation and disease and insect resistance.

5.9. Seed genotyping

NGS and high throughput genotyping will enable large numbers of specific traits to be linked to specific markers, thus increasing the importance of marker assisted selection in the breeding of cotton. However the ability to easily stack a moderate number of traits becomes quickly impractical in the field, where extremely large population would need to be grown to identify plants carrying all the correct combination of markers. To overcome this problem a number of companies have developed non-destructive seed based screening methodologies to screen markers. 'Seed chippers' [96] use complicated robotics and liquid handling equipment to remove small portions of embryo, extracting DNA and performing genotyping. This technology was first demonstrated in maize and then soybean but is beginning to be adapted to work in cotton. Seed-based screening will play an essential role in future transgenic trait stacking and marker-assisted breeding of disease and fibre traits in cotton as it will allow the very small proportion of plants with the favorable allele combination to be identified. Seeds selected that contain the correct marker combinations will be planted in the field and selections for fibre and yield traits performed normally, with the knowledge that other desired traits are already present.

6. Success stories of germplasm utilization in Australian cotton improvement

6.1. Disease resistance

As mentioned previously, bacterial blight was the most important disease in Australian cotton when the modern Industry first started with estimated yield losses up to 20%. Bacterial blight resistance was introduced to the CSIRO breeding program from Tamcot SP37, which supplied the B_2, B_3 and B_7 genes [97]. The first blight resistant cultivar was released in 1985 and had rapid adoption. Importantly, this cultivar also had high yield, high gin turnout, wide adaptation and the okra leaf trait. This cultivar was selected from a large breeding population that

was initially developed in 1974, which demonstrates the considerable lag from introduction of a trait to commercial release.

Verticillium wilt continues to be an import disease for the Australian industry. However, there have been major successes in developing cultivars with good resistance to this disease. A major breakthrough was the release of Sicala V-1 in 1991, the first CSIRO cultivar with significant resistance, and its higher yielding replacement, Sicala V-2 in 1994. These cultivars show greatly reduced levels of infection, less severe symptoms and higher yields compared to their predecessors and have transformed those regions where Verticillium wilt severely limited yields. None of the parents of the population these cultivars were developed from had significant resistance to Verticillium wilt and would generally have been considered susceptible. The source of resistance has therefore been considered to be the result of additive effect of alleles at multiple loci. Molecular techniques may one day be able to determine the source and composition of this resistance.

After Fusarium wilt was first identified in Australian cotton in the early 1990s, new cultivars were required to allow viable production in the regions where the disease was present. It was quickly established that there was cultivar variability for resistance and a few cultivars did show some degree of resistance (Sicot 189, Delta Emerald). Most other commercial cultivars were very susceptible to the disease and had survival rate of close to zero in some situations. An initial screen from the CSIRO germplasm collection in 1994 tested 36 *G. hirsutum* genotypes from the USA, South America, Africa, Asia and Europe. As with the local cultivars, most were susceptible, but a few genotypes showed field survival up to 50% higher than the best Australian cultivars. Over the next few years over 200 genotypes from the collection were evaluated, with material from countries/regions/programs that showed promise in the early evaluations targeted. In 2004 a new cultivar, Sicot F-1, was released from the CSIRO breeding program which had at least twice the resistance of the best cultivar in 1994. Progress in breeding for improved Fusarium resistance continues, with the initial sources of resistance being derived from Indian and Chinese *G. hirsutum* parents and more recently, improved resistance from *G. hirsutum* and *G. barbadense* landrace cottons.

The key to the breeding success of soil borne fungal diseases has been the availability of screening sites with high levels of inoculum and therefore the opportunity for selection of reduced disease incidence and symptoms, together with high yield. The germplasm collections have been invaluable in accessing material for evaluation of resistance.

6.2. Commercial utilisation of the okra leaf mutant

The okra leaf (OL) mutant of upland cotton is characterised by deep lobing and some reduction in leaf area. This L_2^o allele is partially dominant over the normal leaf (l_2) shape. Although considerable research has been done on the effects of this leaf shape, only in Australia have OL cottons been grown on a major scale, accounting for between 40-60% of Australian seed sales between 1987 and 1993. As reviewed by Thomson [6] the published benefits of OL include host plant resistance to a range of insect and mite pests (boll weevil *Anthonomus grandis*, whitefly *Bemisia tabaci*, spider mites *Tetranychus spp.* and bollworms *Heliothis* and *Helicoverpa spp.*), earlier maturity, increased water-use efficiency and reduced boll rot. However, the OL

trait has been associated with increased numbers of plant bugs (*Lygus spp.*), mirids (*Creontiades spp.*) and cotton flea hoppers (*Pseudatomoscelis seriatas*). The yield effect of OL has been reported as being anywhere from negative to positive and is thus far from clear cut.

In the 1970s, the CSIRO breeding program at Narrabri was focused on increasing yield, quality and disease resistance, but also had a strong commitment to increasing resistance to insects. To this end, significant effort was allocated to the OL trait, with large breeding populations developed as well as considerable research into the effect on yield. It was concluded that while there was a small (5%) reduction in yield potential, under commercial production systems with normal levels of insect damage there was no significant yield difference between OL and normal leaf (NL) near-isogenic lines. In addition, in situations where some insect damage was permitted, OL out-yielded NL by up to 10% [6]. A similar observation was reported by Brook *et al.* [98] who stated that OL was infested with fewer pests than NL and suffered less damage from a given density of pests and the reduction in yield caused by pest damage at high yield levels was less in OL than NL.

Thomson [6] reported the reasons why OL has been so successful in Australia when it has either largely failed or been ignored in other countries. As mentioned at the beginning of this chapter, the modern cotton industry is relatively new in Australia, only being a significant industry since the early 1960s. Because growers did not have a long tradition of growing cotton, they had no fixation on what a cotton plant should look like – their concern being solely on performance. The absence of a yield penalty in commercial production systems as mentioned above was also a significant factor in the adoption of OL cottons. Importantly, the Australian OL cottons were not developed by backcross as most other studies have reported (e.g. [99-101]. Rather, as reported by [6], they were developed by hybridisation using wide crosses and large populations with heavy selection pressure for yield performance. Research, together with commercial experience demonstrated the host plant resistance of OL to spider mites and *Helicoverpa* as well as improved insecticide efficacy probably due to better spray penetration in the crop canopy. Lastly and of significant importance, the first OL cotton released for commercial sale (Siokra 1-1) was also the first cultivar in Australia to be completely resistant to bacterial blight. This was undoubtedly a significant factor in its rapid acceptance by growers.

7. Conclusions

The challenge of maintaining genetic diversity in this era of fierce intellectual property protection and commercial reality is significant. One of the ways this is being addressed in the CSIRO breeding program is by developing diversity through utilisation of the genetic resources outside the cultivated tetraploids. This can mean using landrace cottons, but more importantly evaluating traits in the secondary germplasm pool (A and D genomes) and accessing those traits by developing synthetic tetraploids. This approach to diversity requires considerable effort and a long-term vision; however, it does have the potential to pay substantial dividends in unlocking traits previously unavailable from within elite cultivars. To justify this approach, there generally must be present a crucial trait of commercial importance not available in existing germplasm and/or forward-looking funding.

Access to a broad germplasm pool and the ability to import material from other breeding programs combined with rigorous selection regimes for yield and quality under the unique Australian climate has meant that the CSIRO cotton breeding program has maintained yield progress while many other programs are plateauing in yield. Australian cultivars have had considerable success in the US, particularly Texas where there are some similarities in climatic challenges and are now being marketed worldwide under the FiberMax™ brand.

New DNA sequencing technologies will accelerate the identification of genes involved in important agronomic traits, and via marker assisted seed selection, new and novel traits will be introgressed into elite cultivars that should further improve Australian cotton yield and quality.

Author details

Warwick N. Stiller[1*] and Iain W. Wilson[2]

*Address all correspondence to: warwick.stiller@csiro.au

1 CSIRO Plant Industry, Narrabri, NSW, Australia

2 CSIRO Plant Industry, Canberra, ACT, Australia

References

[1] Basinski JJ. The cotton growing industry in Australia. Journal of the Australian Institute of Agricultural Science 1965;31:206-22.

[2] Liu SM, Constable GA, Reid PE, Stiller WN, Cullis BR. The interaction between breeding and crop management in improved cotton yield. Field Crops Research 2013;148:49-60.

[3] Low A. Cotton cultivars in Australia. Cotton Growing Review 1974;51:26-34.

[4] Thomson NJ, Basinski JJ. Cotton in the Ord Valley of northern Australia. Cotton Growing Review 1962;34:81-92.

[5] Thomson NJ. Host plant resistance for cotton. Journal of the Australian Institute of Agricultural Science 1987;53:262-70.

[6] Thomson NJ, editor. Commercial utilisation of the okra leaf mutant of cotton-the Australian experience. Proc World Cotton Research Conf-1, Feb 14-17 1994, Brisbane, Australia: CSIRO; 1994

[7] Constable GA, Reid PE, Thomson NJ. Approaches utilized in breeding and development of cotton cultivars in Australia. In: Jenkins J, Saha S, editors. Genetic Improve-

ment of Cotton: emerging technologies. Enfield, New Hampshire: Science Publishers, Inc.; 2001. p1-15.

[8] Fitt GP, Wilson LJ, Kelly D, Mensah R. Advances in integrated pest management as a component of sustainable agriculture: The case study of the Australian Cotton Industry. In: Peshin R, Dhawan AK, editors. Integrated Pest Management: Innovation-Development Process: Springer Science and Bussiness Media BV; 2009. p507-24.

[9] Wilson LJ, Mensah RK, Fitt GP. Implementing IPM in Australian cotton. In: Horowitz AR, Ishaaya I, editors. Novel approaches to insect pest management in field and protected crops. Netherlands: Springer Press; 2004. p97-118.

[10] Wilson LJ, Bauer LR, Lally DA. Insecticide-induced increases in aphid abundance in cotton. Australian Journal of Entomology 1999;30;38:242-3.

[11] Wilson LJ, Bauer LR, Lally DA. Effect of early season insecticide use on predators and outbreaks of spider mites (Acari : Tetranychidae) in cotton. Bulletin of Entomology Research 1998;88(4):477-88.

[12] Wilson L, Downes S, Khan M, Whitehouse M, Baker G, GRundy P, et al. IPM in the transgenic era: A review of the challenges from emerging pests in Australian cotton systems. Crop and Pasture Science 2013;64:737-49.

[13] Constable G, Llewellyn D, Wilson L, Stiller W. An industry transformed" The impact of GM technology on Australian cotton production. Farm Policy Journal 2011;8(1): 23-41.

[14] Gunning RV, Byrne FJ, Conde BD, Connelly MI, Hergstrom K, Devonshire AL. First Report of B-Biotype *Bemisia-Tabaci* (Gennadius) (Hemiptera, Aleyrodidae) in Australia. Journal of the Australian Entomological Society 1995;34:116-20.

[15] Sequeira RV, Naranjo SE. Sampling and management of *Bemisia tabaci* (Genn.) biotype B in Australian cotton. Crop Protection 2008;27(9):1262-8.

[16] Werth JA, Preston C, Roberts GN, Taylor IN. Weed management practices in glyphosate-tolerant and conventional cotton fields in Australia. Aust Journal of Experimental Agriculture 2006;46(9):1177-83.

[17] Charles GW, Taylor IN. Herbicide resistance and species shift in cotton: the need for an integrated weed management (IWM) approach. In: Swanepoel A, editor. World Cotton Research Conference-3; Cape Town, South Africa: Agricultural Research Council-Institute for Industrial Crops; 2003. p818-28.

[18] Allen SJ. Predominance of race 18 of *Xanthomonas campestris* pv. *malvacearum* on cotton in Australia. Plant Disease 1991;75:43-4.

[19] Schnathorst W, Evans G. Comparative Virulence of American and Australian Isolates of *Verticillium-albo-atrum* in *Gossypium hirsutum*. Plant Disease Reporter 1971;55(11): 977-80.

[20] Allen SJ, editor. The distribution, importance and control of diseases of cotton in Australia. Beltwide Cotton Conference; 1992; Nashville TN: National Cotton Council of America.

[21] Bell AA. Verticillium wilt. In: Hillocks RJ, editor. Cotton Diseases. Wallingford, UK: CAB International; 1992. p87-126.

[22] Kochman JK. Fusarium wilt in cotton-a new record in Australia. Australasian Plant Pathology 1995;24(1):74.

[23] Allen SJ, Lonergan PA. Control strategies for Fusarium wilt of cotton in Australia. In: Dugger P, Richter D, editors. Beltwide Cotton Conference; San Antonio, TX: National Cotton Council of America, Memphis, TN.; 2000.

[24] Kim Y, Hutmacher RB, Davis RM. Characterisation of California isolates of *Fusarium oxysporum* f.sp. *vasinfectum*. Plant Disease 2005;89:366-72.

[25] Colyer PD. A comparison of Fusarium wilt in Australia and the United States. In: Dugger P, Richter D, editors. Beltwide Cotton Conference; Nashville, TN: National Cotton Council of America, Memphis, TN; 2003.

[26] Wang B, Dale ML, Kochman JK. Studies on a pathogenicity assay for screening cotton genrmplasm for resistance to *Fusarium oxysporum* f. sp. *vasinfectum* in the glasshouse. Australian Journal of Experimental Agriculture. 1999;39(8):967-74.

[27] Hillocks RJ. Fusarium wilt. In: Hillocks RJ, editor. Cotton Diseases. Wallingford, UK: CAB International; 1992. p127-60.

[28] Wang B, Brubaker CL, Summerell BA, Thrall PH, Burdon JJ. Local origin of two vegetative compatibility groups of *Fusarium oxysporum* f. sp *vasinfectum* in Australia. Evolutionary Applications 2010;3(5-6):505-24.

[29] Allen SJ. The Australian cotton industry response to the Fusarium wilt problem. In: Dugger P, Richter D, editors. Beltwide Cotton Conferences; Altanta, GA: National Cotton Council of America, Memphis, TN; 2002.

[30] Ellis MH, Silva TF, Stiller WN, Wilson LJ, Vaslin MFS, Sharman M, et al. Identification of a new Polerovirus (family Luteoviridae) associated with cotton bunchy top disease in Australia. Australasian Plant Pathology 2013;42(3):261-9.

[31] Allen SJ. *Thielaviopsis basicola*, a new record on cotton in Australia. Australasian Plant Pathology 1990;19:24-5.

[32] Bauer B, Smith L, Scheikowski L, Lehane J, Cobon J, O'Neil W. Reniform nematode: A new pest of Australian cotton and it's distribution in the Dawson Valley. The Australian Cotton Grower 2013;34(5):28-32.

[33] Bowman DT. Public cotton breeders-do we need them? Journal of Cotton Science 1999;3:139-52.

[34] Clement JD, Constable GA, Stiller WN, Liu SM. Negative associations still exist between yield and fibre quality in cotton breeding programs in Australia and USA. Field Crops Research 2012;128:1-7.

[35] Stokes C, Howden M. Adapting agriculture to climate change. In: Cleugh H, Smith MS, Battaglia M, Graham P, editors. Climate change: science and solutions for Australia. Collingwood, VIC, Australia: CSIRO; 2011. p85-96.

[36] Rochester IJ. Assessing internal crop nitrogen use efficiency in high-yielding irrigated cotton. Nutrient Cycling in Agroecosystems 2011;90(1):147-56.

[37] Rochester IJ. Using seed nitrogen concentration to estimate crop N use-efficiency in high-yielding irrigated cotton. Field Crops Research 2012;127:140-5.

[38] Craven LA, Stewart JM, Brown AHD, Grace JP. The Australian wild species of *Gossypium*. In: Constable GA, Forrester NW, editors. World Cotton Research Conference-1; February 14-17, 1994; Brisbane, Australia: CSIRO, Melbourne; 1994. p278-81.

[39] Stewart JM. Potential for crop improvement with exotic germplasm and genetic engineering. In: Constable G, Forrester N, editors. Challenging the future: Proceedings of the World Cotton Research Conference-1; February 14-17; Brisbane, Australia: CSIRO, Melbourne; 1994. p313-27.

[40] Anon. The biology and ecology of cotton (*Gossypium hirsutum*) in Australia. 2002. August 2002. http://www.ogtr.gov.au/internet/ogtr/publishing.nsf/content/cotton-3/$FILE/biologycotton.pdf

[41] Brown AHD, Brubaker CL. Genetics and the conservation and use of Australian wild relatives of crops. Australian Journal of Botany 2000;48(3):297-303.

[42] Brubaker CL, Brown AHD. The use of multiple alien chromosome addition aneuploids facilitates genetic linkage mapping of the Gossypium G genome. Genome 2003;46(5):774-91.

[43] Lopez-Lavalle LAB, Matheson B, Brubaker CL. A genetic map of an Australian wild Gossypium C genome and assignment of homoeologies with tetraploid cultivated cotton. Genome. 2011;54(9):779-94.

[44] Brubaker CL, Brown AHD, Stewart JM, Kilby MJ, Grace JP. Production of fertile hybrid germplasm with diploid Australian Gossypium species for cotton improvement. Euphytica 1999;108(3):199-213.

[45] Mergeai G. Forty years of genetic improvement of cotton through interspecific hybridisation at Gembloux Agricultural University: Achievements and prospects. In: Swanepoel A, editor. World Cotton Research Conference-3; Cape Town, South Africa: Agricultural Research Council-Institute for Industrial Crops; 2003. p119-33.

[46] Lopez-Lavalle LAB, McFadden H, Brubaker CL. The effect of Gossypium C-genome chromosomes on resistance to Fusarium wilt in allotetraploid cotton. Theoretical Applied Genetics 2007;115(4):477-88.

[47] McFadden H, Beasley D, Brubaker CL. Assessment of Gossypium sturtianum and G-australe as potential sources of Fusarium wilt resistance to cotton. Euphytica 2004;138(1):61-72.

[48] Brubaker CL, Benson CG, Miller C, Leach DN. Occurrence of terpenoid aldehydes and lysigenous cavities in the 'glandless' seeds of Australian Gossypium species. Australian Journal of Botany 1996;44(5):601-12.

[49] Llewellyn D, Fitt G. Pollen dispersal from two field trials of transgenic cotton in the Namoi Valley, Australia. Molecular Breeding 1996;2(2):157-66.

[50] Llewellyn D, Tyson C, Constable G, Duggan B, Beale S, Steel P. Containment of regulated genetically modified cotton in the field. Agriculture, Ecosystems & Environment 2007;121(4):419-29.

[51] Lopez-Lavalle LAB, Gillespie VJ, Tate WA, Ellis MH, Stiller WN, Llewellyn DL, et al. Molecular mapping of a new source of Fusarium wilt resistance in tetraploid cotton (*Gossypium hirsutum* L.). Molecular Breeding 2012;30(2):1181-91.

[52] Lopez-Lavalle LAB, Potter N, Brubaker CL. Development of a rapid, accurate glasshouse bioassay for assessing Fusarium wilt disease responses in cultivated Gossypium species. Plant Pathology 2012;61(6):1112-20.

[53] Lacape JM, Jacobs J, Arioli T, Derijcker R, Forestier-Chiron N, Llewellyn D, et al. A new interspecific, *Gossypium hirsutum* x *G. barbadense,* RIL population: towards a unified consensus linkage map of tetraploid cotton. Theoretical Applied Genetics 2009;119(2):281-92.

[54] Lacape JM, Llewellyn D, Jacobs J, Arioli T, Becker D, Calhoun S, et al. Meta-analysis of cotton fiber quality QTLs across diverse environments in a *Gossypium hirsutum* x *G. barbadense* RIL population. BMC Plant Biology 2010;28;10.

[55] Lacape JM, Gawrysiak G, Cao TV, Viot C, Llewellyn D, Liu SM, et al. Mapping QTLs for traits related to phenology, morphology and yield components in an inter-specific *Gossypium hirsutum* x *G. barbadense* cotton RIL population. Field Crops Research 2013;144:256-67.

[56] Liu SM, Llewellyn DJ, Stiller WN, Jacobs J, Lacape JM, Constable GA. Heritability and predicted selection response of yield components and fibre properties in an inter-specific derived RIL population of cotton. Euphytica 2011;178(3):309-20.

[57] Thomson NJ, Reid PE, Williams ER. Effects of the Okra Leaf, Nectariless, Frego Bract and Glabrous Conditions on Yield and Quality of Cotton Lines. Euphytica 1987;36(2): 545-53.

[58] Wu YR, Machado AC, White RG, Llewellyn DJ, Dennis ES. Expression profiling identifies genes expressed early during lint fibre initiation in cotton. Plant Cell Physiology 2006;47(1):107-27.

[59] Brown N, Smith CW, Auld D, Hequet EF. Improvement of Upland Cotton Fiber Quality through Mutation of TAM 94L-25. Crop Science 2013;53(2):452-9.

[60] Bechere E, Auld DL, Dotray PA, Gilbert LV, Kebede H. Imazamox Tolerance in Mutation-Derived Lines of Upland Cotton. Crop Science 2009;49(5):1586-92.

[61] Bechere E, Auld DL, Hequet E. Development of 'naked-tufted' seed coat mutants for potential use in cotton production. Euphytica 2009;167(3):333-9.

[62] Herring AD, Auld DL, Ethridge MD, Hequet EF, Bechere E, Green CJ, et al. Inheritance of fiber quality and lint yield in a chemically mutated population of cotton. Euphytica 2004;136(3):333-9.

[63] Cousins YL, Lyon BR, Llewellyn DJ. Transformation of an Australian Cotton Cultivar-Prospects for Cotton Improvement through Genetic-Engineering. Australian Journal of Plant Physiology 1991;18(5):481-94.

[64] Machado A, Wu YR, Yang YM, Llewellyn DJ, Dennis ES. The MYB transcription factor GhMYB25 regulates early fibre and trichome development. Plant Journal 2009;59(1):52-62.

[65] Walford SA, Wu YR, Llewellyn DJ, Dennis ES. GhMYB25-like: a key factor in early cotton fibre development. Plant Journal 2011;65(5):785-97.

[66] Walford SA, Wu YR, Llewellyn DJ, Dennis ES. Epidermal cell differentiation in cotton mediated by the homeodomain leucine zipper gene, GhHD-1. Plant Journal 2012;71(3):464-78.

[67] Pu L, Li Q, Fan XP, Yang WC, Xue YB. The R2R3 MYB Transcription Factor GhMYB109 Is Required for Cotton Fiber Development. Genetics 2008180(2):811-20.

[68] Ruan YL, Llewellyn DJ, Furbank RT. Suppression of sucrose synthase gene expression represses cotton fiber cell initiation, elongation, and seed development. Plant Cell 2003;15(4):952-64.

[69] Xu SM, Brill E, Llewellyn DJ, Furbank RT, Ruan YL. Overexpression of a Potato Sucrose Synthase Gene in Cotton Accelerates Leaf Expansion, Reduces Seed Abortion, and Enhances Fiber Production. Molecular Plant 2012;5(2):430-41.

[70] Murray F, Llewellyn D, McFadden H, Last D, Dennis ES, Peacock WJ. Expression of the *Talaromyces flavus* glucose oxidase gene in cotton and tobacco reduces fungal infection, but is also phytotoxic. Molecular Breeding 1999;5(3):219-32.

[71] Townsend BJ, Poole A, Blake CJ, Llewellyn DJ. Antisense suppression of a (+)-delta-cadinene synthase gene in cotton prevents the induction of this defense response

gene during bacterial blight infection but not its constitutive expression. Plant Physiology 2005;138(1):516-28.

[72] Liu Q, Singh SP, Green AG. High-stearic and high-oleic cottonseed oils produced by hairpin RNA-mediated post-transcriptional gene silencing. Plant Physiology 2002;129(4):1732-43.

[73] Townsend BJ, Llewellyn DJ. Spatial and temporal regulation of a soybean (Glycine max) lectin promoter in transgenic cotton (*Gossypium hirsutum*). Functional Plant Biology. 2002;29(7):835-43.

[74] Schunmann PHD, Llewellyn DJ, Surin B, Boevink P, De Feyter RC, Waterhouse PM. A suite of novel promoters and terminators for plant biotechnology. Functional Plant Biology. 2003;30(4):443-52.

[75] Amarasinghe B, Faivre N, Wu, Udall J, Dennis E, Constable G, et al. Genomic approaches to the discovery of promoters for sustained expression in cotton (*Gossypium hirsutum* L.) under field conditions: expression analysis in transgenic cotton and Arabidopsis of a Rubisco small subunit promoter identified using EST sequence analysis and cDNA microarrays. Plant Biotechnology 2006;23:437-50.

[76] Abdalla AM, Reddy OUK, El-Zik KM, Pepper AE. Genetic diversity and relationships of diploid and tetraploid cottons revealed using AFLP. Theoretical Applied Genetics 2001;102(2-3):222-9.

[77] Wendel JF, Brubaker CL, Percival AE. Genetic Diversity in *Gossypium hirsutum* and the Origin of Upland Cotton. American Journal of Botany 1992;79(11):1291-310.

[78] Iqbal MJ, Aziz N, Saeed NA, Zafar Y, Malik KA. Genetic diversity evaluation of some elite cotton varieties by RAPD analysis. Theoretical Applied Genetics 1997;94(1): 139-44.

[79] Wendel JF, Olson PD, Stewart JM. Genetic Diversity, Introgression, and Independent Domestication of Old-World Cultivated Cottons. American Journal of Botany 1989;76(12):1795-806.

[80] Wendel JF, Cronn RC. Polyploidy and the evolutionary history of cotton. Advances in Agronomy, Vol 78. 2003;78:139-86.

[81] Schuster SC. Next-generation sequencing transforms today's biology. Nature Methods 2008;5(1):16-8.

[82] Paterson AH, Wendel JF, Gundlach H, Guo H, Jenkins J, Jin DC, et al. Repeated polyploidization of Gossypium genomes and the evolution of spinnable cotton fibres. Nature 2012;492(7429):423-+.

[83] Huang XH, Wei XH, Sang T, Zhao QA, Feng Q, Zhao Y, et al. Genome-wide association studies of 14 agronomic traits in rice landraces. Nature Genetics 2010;42(11):961-U76.

[84] Cavanagh CR, Chao SM, Wang SC, Huang BE, Stephen S, Kiani S, et al. Genome-wide comparative diversity uncovers multiple targets of selection for improvement in hexaploid wheat landraces and cultivars. Proceeding of the National Academy of Science U.S.A. 2013;110(20):8057-62.

[85] Barbazuk WB, Emrich SJ, Chen HD, Li L, Schnable PS. SNP discovery via 454 transcriptome sequencing. Plant Journal 2007;51(5):910-8.

[86] Deschamps S, Llaca V, May GD. Genotyping-by-Sequencing in Plants. Biology 2012;1(3):460-83.

[87] Gore MA, Chia JM, Elshire RJ, Sun Q, Ersoz ES, Hurwitz BL, et al. A First-Generation Haplotype Map of Maize. Science 2009;326(5956):1115-7.

[88] Fang DD, Hinze LL, Percy RG, Li P, Deng D, Thyssen G. A microsatellite-based genome-wide analysis of genetic diversity and linkage disequilibrium in Upland cotton (*Gossypium hirsutum* L.) cultivars from major cotton-growing countries. Euphytica 2013;191(3):391-401.

[89] de Bakker PIW, Yelensky R, Pe'er I, Gabriel SB, Daly MJ, Altshuler D. Efficiency and power in genetic association studies. Nature Genetics 2005;37(11):1217-23.

[90] Gupta PK, Rustgi S, Kulwal PL. Linkage disequilibrium and association studies in higher plants: Present status and future prospects. Plant Molecular Biology 2005;57(4):461-85.

[91] Abdurakhmonov IY, Kohel RJ, Yu JZ, Pepper AE, Abdullaev AA, Kushanov FN, et al. Molecular diversity and association mapping of fiber quality traits in exotic *G. hirsutum* L. germplasm. Genomics 2008;92(6):478-87.

[92] Abdurakhmonov IY, Saha S, Jenkins JN, Buriev ZT, Shermatov SE, Scheffler BE, et al. Linkage disequilibrium based association mapping of fiber quality traits in *G. hirsutum* L. variety germplasm. Genetica 2009;136(3):401-17.

[93] Neves LG, Davis JM, Barbazuk WB, Kirst M. Whole-exome targeted sequencing of the uncharacterized pine genome. Plant Journal 2013;75(1):146-56.

[94] Hodges E, Xuan Z, Balija V, Kramer M, Molla MN, Smith SW, et al. Genome-wide in situ exon capture for selective resequencing. Nature Genetics 2007;39(12):1522-7.

[95] Wang TL, Uauy C, Robson F, Till B. TILLING in extremis. Plant Biotechnology Journal 2012;10(7):761-72.

[96] Gao SB, Martinez C, Skinner DJ, Krivanek AF, Crouch JH, Xu YB. Development of a seed DNA-based genotyping system for marker-assisted selection in maize. Molecular Breeding 2008;22(3):477-94.

[97] Bird LS. Registration of Tamcot SP37H cotton. Crop Science 1979;19:412.

[98] Brook KD, Hearn AB, Kelly CF. Response of Cotton to Damage by Insect Pests in Australia-Compensation for Early Season Fruit Damage. Journal of Economic Entomology 1992;85(4):1378-86.

[99] Andries JA, Jones JE, Sloane LW, Marshall JG. Effects of okra leaf shape on boll rot, yield, and other important characters of upland cotton, *Gossypium hirsutum* l. Crop Science 1969;9:705-10.

[100] Meredith WR. Effect of environments and genetic backgrounds on evaluation of cotton isolines. Crop Science 1983;23:51-4.

[101] Meredith WR, Wells R. Normal vs. Okra leaf yield interactions in cotton. I.Performance of near isogenic lines from bulk populations. Crop Science 1986;26:219-22.