

Chapter 11

Diversity Arrays Technology (DArT) Markers for Genetic Diversity

Dariusz Grzebelus

Abstract Diversity Arrays Technology (DArT) is a hybridization-based high-throughput genotyping technology that was proposed in the beginning of the twenty-first century as an efficient and cost-effective alternative to existing genotyping systems, and since then it has been used extensively to explore genetic diversity in many plant species. In this chapter, we describe the principles behind DArT genotyping, summarize the research on plant genetic diversity utilizing the DArT system, discuss advantages, limitations, and perspectives of the technology.

Keywords Genetic diversity · Genotyping · High-throughput · Cost-effective · DArT system · Genetic erosion · Crop domestication

11.1 Introduction

Numerous marker technologies facilitating studies and management of plant genetic diversity have been developed over the past few decades. On one hand, marker-based strategies help in investigating species diversity, genetic erosion, crop domestication, etc. On the other hand, they are widely used in crop improvement, allowing more effective utilization of genetic diversity. Historically, the most popular systems were (1) restriction fragment length polymorphisms (RFLP, Botstein et al. 1980), (2) randomly amplified polymorphic DNA (RAPD, Williams et al. 1990), (3) amplified fragment length polymorphism (AFLP, Vos et al. 1995), (4) microsatellites (simple sequence repeats; SSR, Powell et al. 1996), and (5) single nucleotide polymorphisms (SNP, Rafalski 2002). They have been extensively used to genotype plants. RAPD and AFLP systems do not require any prior information on the sequence of

D. Grzebelus (✉)

Institute of Plant Biology and Biotechnology, University of Agriculture in Krakow,
Al. 29 Listopada 54, 31-425 Krakow, Poland
e-mail: d.grzebelus@ogr.ur.krakow.pl

polymorphic sites. They were widely used in the last decade of the twentieth century, as no cost-efficient DNA sequencing technologies were available at that time. In contrast, development of SSR, and particularly SNP markers, requires that sequences of polymorphic sites are known, which allowed their wider introduction only after data from numerous plant genome and transcriptome sequencing projects begun to accumulate in the past decade (Zalapa et al. 2012). Nevertheless, for many crops financial resources are still too low to initiate NGS-based marker discovery and arbitrary markers remain an option of interest for investigation of species genetic diversity.

The genotyping systems described above varied with respect to their capability for rapid identification of large numbers of markers. Most systems provided low- to medium-throughput efficiency, as they relied on sequential identification of polymorphisms, typically by means of agarose or polyacrylamide electrophoresis. Only SNP markers can be identified with several commercially available high-throughput genotyping platforms (reviewed by Gupta et al. 2008). Diversity Arrays Technology (DArT) markers provided a unique option of cost-efficient parallel genotyping with a set of hundreds to thousands of arbitrary markers in a single assay utilizing microarrays. By scoring presence or absence of arbitrary restriction fragments in genomic representations, DArT produces reproducible whole-genome fingerprints (Jaccoud et al. 2001). Here, we describe the principles of the DArT system and present an overview of its applications for assessment of genetic diversity in plants.

11.2 Principles of the Diversity Arrays Technology

Diversity Array Technology (DArT) is a microarray-based molecular marker system allowing cost-efficient (per data point) high-throughput genotyping of any organism. It was developed as a hybridization-based alternative to existing genotyping technologies. Importantly, DArT genotyping does not require any prior knowledge of the genome sequence (Jaccoud et al. 2001). It has been widely applied in plant science and proven to perform well for many species (Kilian et al. 2005). An updated list of reports using DArT markers for evaluating genetic diversity in plants is shown in Table 11.1.

Generally, 100 ng of genomic DNA is enough to genotype more than 7000 genomic loci in parallel in a single-reaction assay. DArT markers are strictly biallelic and are usually scored as presence versus absence variants, where the 'present' state is dominant over the 'absent' state. However, they may also be scored as hemi-dominant taking into account signal intensity as a reflection of the dosage effect (double dose vs. single dose vs. absence). The observed polymorphisms usually result from single nucleotide substitutions within restriction sites or InDels including restriction sites, but they can also be caused by differences in the methylation status (see below). Nevertheless, the structural polymorphisms account for more than 90 % of the identified variability (Wittenberg et al. 2005) and are inherited in a simple Mendelian fashion.

Table 11.1 Research projects on plant genetic diversity utilizing the Diversity Arrays Technology (DArT) platform

Species	Objective of the study	No. of DArT markers	Reference
<i>Aegilops tauschii</i>	Population structure and diversity	4449	Sohail et al. (2012)
<i>Asplenium viride</i>	Evolution of substrate specificity	444	James et al. (2008)
Banana (<i>Musa</i> sp.)	Fingerprinting and diversity	836	Risterucci et al. (2009)
Banana (<i>Musa</i> sp.)	Genetic diversity	653	Amorim et al. (2009)
Barley (<i>Hordeum vulgare</i>)	Genetic variability assessment	271	Ovesná et al. (2013)
Carrot (<i>Daucus carota</i>)	Genetic diversity, wild versus cultivated	900	Grzebelus et al. (2014)
Cassava (<i>Manihot esculenta</i>)	Genetic diversity	1000	Xia et al. (2005), Hurtado et al. (2008)
Common beans (<i>Phaseolus vulgaris</i>)	Germplasm genetic diversity	2501	Briñez et al. (2012)
<i>Eucalyptus grandis</i>	Fingerprinting	104	Lezar et al. (2004)
<i>Garovagliaelegans</i>	Phylogeography	905	James et al. (2008)
Hop (<i>Humulus lupulus</i>)	Fingerprinting and diversity	730	Howard et al. (2011)
Lesquerella (<i>Physaria</i> sp.)	Genetic diversity	2833	Cruz et al. (2013)
Mungbean (<i>Vigna radiata</i>)	Genetic diversity	1125	Hang Vu et al. (2012)
Oat (<i>Avena sativa</i>)	Genetic diversity	1295	Tinker et al. (2009)
Olive (<i>Olea europaea</i>)	Genetic diversity	2031	Domínguez-García et al. (2012); Atienza et al. (2013)
Pigeonpea (<i>Cajanus cajan</i>)	Genetic diversity, wild versus cultivated	700	Yang et al. (2006)
Rapeseed (<i>Brassica napus</i>)	Genetic diversity	1547	Raman et al. (2012)
Rice (<i>Oryza sativa</i>)	Fingerprinting and diversity	1152	Xie et al. (2006)
Rye (<i>Secale cereale</i>)	Genetic diversity	1022	Bolibok-Bragoszewska et al. (2009)
<i>Sorghum bicolor</i>	Genetic diversity	508	Mace et al. (2008)
Sugarcane (<i>Saccharum</i> sp.)	Genetic diversity	667	Heller-Uszynska et al. (2011)
Tall fescue (<i>Festuca arundinacea</i>)	Genetic diversity	190	Baird et al. (2012)
<i>Triticum monococcum</i>	Genetic diversity	846	Jing et al. (2009)
Wheat (<i>Triticum aestivum</i>)	Genetic diversity	411	Akbari et al. (2006), White et al. (2008)

(continued)

Table 11.1 (continued)

Species	Objective of the study	No. of DArT markers	Reference
Wheat (<i>Triticum aestivum</i>)	Population structure and association mapping	318	Crossa et al. (2007)
Wheat (<i>Triticum aestivum</i>)	Genetic diversity, resistance to fusarium head blight	409	Badea et al. (2008)
Wild potatoes (<i>Solanum bulbocastanum</i> , <i>S. commersonii</i>)	Inter-species diversity	1423	Traini et al. (2013)

11.2.1 The DArT System

The first step in the development of the DArT genotyping platform for a species of interest is the assembly of a set of arbitrary genomic DNA fragments representative of the germplasm under investigation using a procedure called ‘complexity reduction’ (Fig. 11.1). The fragments are derived from a collection of individuals representing the primary gene pool of the species. A few complexity reduction strategies have been applied by different authors. Here we present the most widely implemented strategy in which the fragments are obtained by double restriction digestion of pooled genomic DNAs of plants comprising the collection with *Pst*I (6-cutter) and a frequently cutting restriction enzyme (4-cutter, e.g. *Taq*I, *Bst*NI, *Apo*I, etc.). *Pst*I is used because of its methylation sensitivity—it does not cut in methylated regions and thus it allows getting rid of the heavily methylated highly repetitive fraction of the genome. It is essential to carefully select the most suitable frequently cutting restriction enzyme, as it was shown that their ability to reveal polymorphisms may differ significantly, especially in larger genomes comprising more repetitive DNA (Wenzl et al. 2004). Subsequently, *Pst*I- and 4-cutter-restriction site-specific adaptors are ligated to the ends of the restriction fragments and adaptor-specific primers are used to amplify them.

Additional modifications of complexity reduction methods, used mostly for analyses of more complex genomes, include the use of fragments developed from amplification of regions adjacent to insertion sites of miniature inverted-repeat transposable elements (MITEs) and application of suppression subtractive hybridization (SSH, Diatchenko et al. 1996) to enrich genomic representations with polymorphic clones (James et al. 2008; Mace et al. 2008; Heller-Uszynska 2011).

The amplicons are then ligated into a plasmid and cloned in *Escherichia coli*. Individual *E. coli* colonies carrying inserts are arrayed on 384-well plates. The set of inserts comprising the library is called ‘genomic representation’ and can be characterized by the level of complexity depending on the size of the studied genome, number of fragments in the library and size of fragments, which usually is in the range of 300–700 bp. Typically, the genomic representation of a plant genome contains no more than a few percent of the whole genome. The library

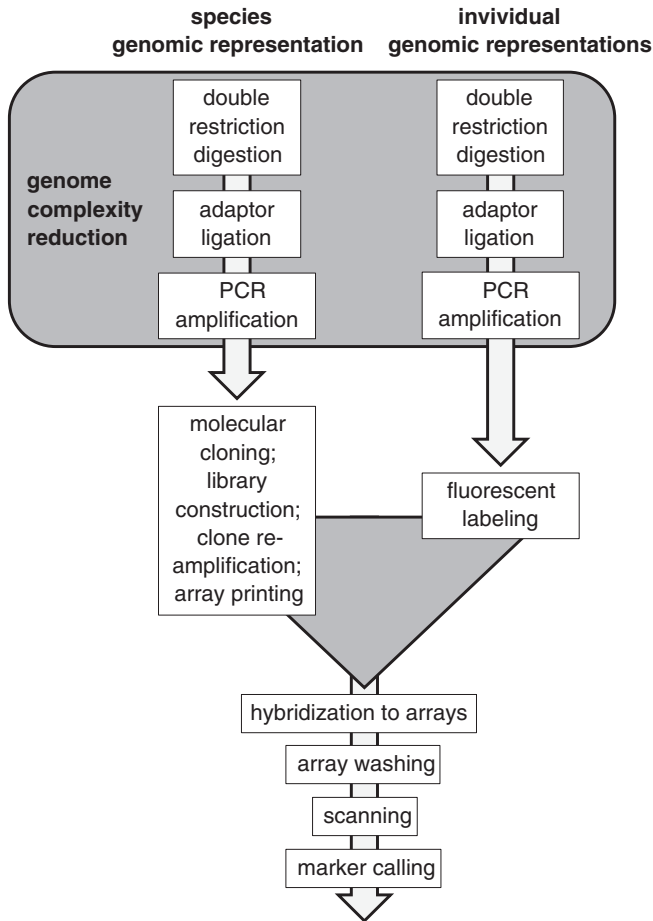


Fig. 11.1 A schematic diagram presenting steps of the Diversity Array Technology genotyping platform

is used to prepare spotted glass microarrays for routine assays. For this purpose, inserts are reamplified from plasmids using a pair of universal vector-specific primers, so that each amplicon carries a segment derived from the genomic DNA of interest and vector segments adjacent to the multiple cloning site, the latter being present in all spotted DNA fragments.

Genomic representations of individuals subject to genotyping (called ‘targets’) are obtained from single genomic DNA isolations using the above strategy (Fig. 11.1). They are fluorescently labeled and hybridized to the glass microarrays on which the genomic representation of the species was spotted. A multiple cloning site of the vector (called ‘reference’) fluorescently labeled with a dye different from that used for the genomic representation is also used for hybridization, in parallel with the target. The reference provides quality control for each spot

as it allows measurement of the signal-to-noise ratio. Following hybridization, the microarrays are washed, scanned with a confocal laser scanner, and analyzed with a dedicated software called DArT soft, performing image analysis, marker discovery, and marker scoring (Kilian et al. 2005).

11.2.2 Limitations of DArT Markers

Three major issues, i.e., low level of polymorphism, redundancy, and sensitivity to methylation, may affect optimal implementation of the DArT genotyping platform. Typically, only from 5 to 30 % of all spotted fragments allow identification of polymorphisms. In order to making the DArT genotyping more effective, it is possible to rearrange the initial array to remove all nonpolymorphic and unreliable clones. In a number of more advanced DArT genotyping programs, a strategy is used which involves initial development of ‘discovery arrays’, identification of the most informative DArT markers, subsequent re-arraying, and assembly of a final ‘genotyping array’ (Gupta et al. 2008).

Redundancy is caused by the presence of multiple clones in the genomic representation library that were derived from the same genomic region. Grzebelus et al. (2014) estimated that a very high fraction of DArT clones, reaching 50 %, were redundant in the carrot discovery array, while only 11 and 16 % redundancy was reported in *Asplenium* and *Garovaglia* arrays, respectively (James et al. 2008). There are two possible causes of the observed redundancy; (1) the redundant fragments originated from repetitive regions and (2) the redundant fragments were preferentially PCR-amplified. While the presence of repetitive fragments can be limited by careful selection of the combination of restriction enzymes, the amplification issues can at least in part be solved by optimization of cycling parameters, including primer annealing temperatures and limiting the number of PCR cycles.

As *Pst*I restriction enzyme routinely used for preparation of genomic representations is methylation-sensitive, a fraction of observed polymorphisms can originate from different methylation status of the same sequence. It was reported that for less than 10 % of DArT markers in *Arabidopsis* no sequence polymorphism could have been detected, implying that they represented methylation variants (Wittenberg et al. 2005). Interestingly, at least one of DArT markers showing strong signature for selection in the cultivated carrot was apparently a result of a systematic difference of the methylation status rather than sequence variability (D. Grzebelus, unpublished). Thus, even if sensitivity to methylation is generally undesired, in particular cases it can be viewed as an additional advantage of the technology, depending on the research objectives.

Bolibok-Bragoszewska et al. (2009) stressed the fact that the dominant character of DArT markers may limit their usefulness for the assessment of genetic diversity in highly heterozygous obligatory outcrossing species. However, other authors postulated that the high number of DArT markers identified per assay combined with the use of the most appropriate strategy for inferring population

structure provided satisfactory results. Also, it is possible to score DArT markers in a hemi-dominant (dosage-dependent) manner to identify the heterozygote state (Kilian et al. 2005).

11.3 Application of the DArT Marker System for Evaluation of Genetic Diversity

The technology was originally developed for rice, a diploid crop with a small genome of 430 Mb. In the proof-of-concept paper presenting capability of the DArT system to capture genetic variability, Jaccoud et al. (2001) demonstrated that it could be used to investigate genetic diversity of rice cultivars of different origin. Xia et al. (2006) developed a general purpose rice DArT platform and used it to study genetic diversity in 24 rice cultivars originating from the Yunnan province, concluding that the level of genetic diversity in rice hybrid cultivars was low, while it was higher in a set of investigated landraces. Recently, Courtois et al. (2013) developed a japonica rice genotyping panel employing an NGS-based variant of DArT called DArT seq (see Perspectives section) and used it to analyze 167 accessions of *O. sativa* var. *japonica* with the purpose of association mapping of root traits. With respect to genetic diversity, they revealed diversity structure comprising six subpopulations, reflecting geographic origin and breeding history. A large number of admixed accessions confirmed gene exchange among subpopulations.

DArT has been extensively used to study genetic diversity in other cereal crops. Ovesná et al. (2013) analyzed genetic diversity in 94 Czech malting barley cultivars. They reported that the level of genetic diversity remained roughly unchanged, but significant shifts in allelic frequency occurred over time, likely resulting from the impact of breeding practices. Old barley cultivars grouped separately from the remaining accessions. As the DArT similarity matrices correlated well with similarity matrices based on agronomical and chemical data, the authors concluded that the DArT method accurately reflected the genetic basis of traits of the investigated barley cultivars. Thirty-one varieties and breeding lines were used to evaluate genetic diversity in rye (*Secale cereale*). All varieties clustered together, while more diversity was observed among breeding lines (Bolibok-Bragoszewska et al. 2009). Mace et al. (2008) developed a DArT platform to investigate genetic diversity in sorghum (*Sorghum bicolor*). They analyzed 90 accessions representing a significant portion of genetic variation in sorghum and showed that they were well separated upon DArT genotyping. Thirteen main clusters were revealed, reflecting the race and origin of accessions grouped in the clusters, as well as their status as B (maintainer female) or R (male parental restorer). Research on wheat and oat is outlined in the section devoted to polyploid species.

One of the early projects aiming at the development of a microarray-based platform was carried out in eucalyptus (Lezar et al. 2004). Twenty-three *Eucalyptus grandis* trees were fingerprinted with a set of 384 arbitrary clones, of which 104

identified polymorphisms. Seventeen full-sib trees could be unequivocally identified on the basis of the assay.

Xia et al. (2005) developed and validated a DArT platform for cassava (*Manihot esculenta*) and investigated genetic diversity among 38 accessions, including wild relatives. It successfully revealed genetic diversity and separated wild accessions from cultivars. Subsequently, Hurtado et al. (2008) used the above-described cassava DArT array to analyze genetic diversity of 436 cassava accessions of African and Latin American origin. While the separation of groups of accessions originating from different continents was revealed with 251 DArT polymorphisms, the expected within-continent genetic diversity could not have been precisely defined.

Several projects on genetic diversity utilizing DArT markers were carried out in legumes. Hang Vu et al. (2012) developed DArT platforms for soybean (*Glycine max*) and mungbean (*Vigna radiata*). The mungbean array was used to elucidate genetic relationships within the genus *Vigna*. Eleven *Vigna* accession were grouped into three clusters, corresponding with *Vigna* sub-genera. Interestingly, a possibility of marker transferability between the *Vigna*- and *Glycine*-specific arrays was reported, allowing their potential use for comparative genomic studies. Briñez et al. (2012) used the DArT system to study genetic diversity in 89 accessions of common beans (*Phaseolus vulgaris*). The two major gene pools of common beans were distinguished and the accessions were classified as either Andean or Mesoamerican.

Application of DArT markers allowed differentiation of 92 hop accessions into two genetically differentiated groups comprising European and North American accessions and a separate group of hybrid cultivars derived from crossings between representatives of the former two groups. Genetic diversity in both geographic groups was similar, while the hybrids showed greater diversity (Howard et al. 2011).

Risterucci et al. (2009) developed a DArT platform for two *Musa* species, *Musa acuminata* and *Musa balbisiana*, donors of A and B genomes, respectively, for cultivated sweet and cooking bananas, most of which are triploids. They analyzed a panel of 168 genotypes and found clear differentiation between the two genomes with further differentiation of *M. acuminata* into two groups, one including mostly wild and the other—mostly cultivated accessions. Grouping of the triploid cultivated forms depended on their constitution; separate groups comprising AAA, AAB, and ABB genomes were revealed. Sub-clusters representing breeding histories and geographic origin were also observed. In another study using DArT markers in *Musa*, Amorim et al. (2009) investigated genetic diversity in a group of 42 carotenoid-rich diploid, triploid, and tetraploid banana accessions. They were divided into two major clusters which did not differentiate diploid and polyploid accessions. Also, no relationship between grouping and carotenoid content was observed.

Domínguez-García et al. (2012) used a collection of 87 olive (*Olea europaea*) accessions representing genetic diversity of the species to develop a DArT platform. In order to validate the array they evaluated genetic diversity in a subset

of 62 accessions, and subsequently Atienza et al. (2013) used the same tool for a large-scale study comprising 323 olive cultivars. Both studies showed the utility of the DArT platform for fingerprinting olive genetic resources. High level of genetic diversity in olive genetic resources is revealed and several duplicated accessions were identified. It was possible to use the olive array to analyze genetic diversity in 42 accessions of wild olive.

Following development of a DArT platform constructed from 107 accessions of *Brassica napus* var. *oleifera* and *Brassica rapa*, Raman et al. (2012) investigated genetic diversity in 89 accessions of rapeseed and 32 accessions of other diploid and tetraploid brassicas, i.e., *B. rapa* (AA), *Brassica juncea* (AABB), and *Brassica carinata* (BBCC). Rapeseed cultivars of the same origin or pedigree tended to form separate groupings within three main clusters. The array was also useful for differentiating species, separating also winter and spring types in the *B. napus* cluster.

A DArT array for carrot was developed by Grzebelus et al. (2014) and used to evaluate genetic diversity in a collection of wild and cultivated accessions of *Daucus carota*. Three major clusters were differentiated, grouping wild, Eastern cultivated, and Western cultivated accessions, which reflected domestication and breeding history of the species. In addition, a subset of DArT markers showing signatures for selection upon domestication was identified.

11.3.1 Performance of the DArT System in Complex Polyploid genomes

The presence of multiple copies of genes in polyploids is prohibitive for many genotyping systems. It was shown that DArT markers can efficiently genotype large polyploid species. DArT markers were effectively applied to genotype the 16Gb hexaploid genome of bread wheat and to analyze intraspecific diversity in *Triticum aestivum* (Akbari et al. 2006). Two separate groupings of European and Australian cultivars were observed in a collection of 62 wheat cultivars, the latter groups being more diverse and having a broader range of adaptation. Crossa et al. (2007) used the wheat DArT array developed by Akbari et al. (2006) to study associations with several traits of agronomic importance. They used two collections of 76 and 94 accessions and revealed a fine population structure of 17 and 15 subpopulations, respectively. The research allowed identification of many new chromosome regions for disease resistance and grain yield in the wheat genome. Badea et al. (2008) evaluated a collection of 87 spring and winter wheat accessions for diversity with respect to resistance to fusarium head blight. They identified six clusters which generally agreed with the origin, growth habit, and pedigree of the studied accessions. White et al. (2008) performed a detailed analysis of spatial and temporal changes of genetic diversity in a collection of 240 wheat varieties of UK, US, and Australian origin. The country of origin accounted for ca. 20 % of the total variation revealed by DArT markers. The highest diversity was observed

in the Australian subset, while the lowest was reported for the UK subset. The D genome occurred to be slightly less diverse than the A and B genomes. Moreover, an upward trend in diversity in the US was noticed, while diversity in Australian and UK varieties remained relatively constant.

Genetic diversity in oat (*Avena sativa*) was analyzed with a set of 182 accessions collected worldwide. Two major groups were observed, comprising spring and winter cultivars, while a finer structure of genetic diversity was attributed to geographic origin and breeding history, with subgroups related to known pedigree structure (Tinker et al. 2009). Baird et al. (2012) investigated genetic diversity in another allohexaploid species of Poaceae, tall fescue (*Festuca arundinacea*). By comparing 97 accessions of turf-type tall fescue with 14 accessions of forage type they concluded that genetic diversity in the turf type was very low and should urgently be broadened.

The DArT system was used to study diversity in sugarcane (*Saccharum* sp.) carrying a polyploid, very complex, and particularly challenging genome. The investigation of 16 genotypes of different pedigree and two modifications of the complexity reduction method revealed high genetic differentiation of sugarcane. The ancestral species of *Saccharum spontaneum* and *Saccharum officinarum* were separated from the rest of the samples (Heller-Uszynska et al. 2011).

11.3.2 Applications of the DArT System to Minor Crops, Wild Crop Relatives and Wild Species

The fact that the DArT platform, unlike other high-throughput genotyping technologies, does not rely on any prior sequence information, facilitates its use in species of little or no agronomic importance. Research on lesquerella (*Physaria* spp.), an alternative oil crop, is an example of the successful application of DArT markers for investigating genetic diversity in the group of novel crops. The DArT platform allowed differentiation of 89 accessions with respect to their species, geographic origin, and breeding status. It also revealed that a substantial genetic diversity was present in *Physaria fendleri* from which several breeding lines have been produced and could be commercialized (Cruz et al. 2013).

Pigeon pea (*Cajanus cajan*) is a representative of a group called ‘orphan’ crops, i.e., a domesticated species of low economic value and limited financial resources allocated to its breeding and conservation, which requires careful calculation of ‘per data point’ genotyping costs. Yang et al. (2006) developed a DArT platform for pigeon pea and evaluated genetic diversity in a set of 232 accessions of *C. cajan* and its wild relatives. Genetic diversity among the cultivated accessions was very low, with only 64 of nearly 700 markers being polymorphic in the cultivated germplasm, indicating a very narrow genetic base. No clear genetic diversity structure was observed in the cultivated group. In contrast, higher diversity was revealed in the group of wild accessions which were grouped according to the

species. The authors concluded that the DArT system is an inexpensive genome profiling technology that is likely to contribute significantly to the effective utilization of genetic diversity in ‘orphan’ crops, such as pigeon pea.

Studies on wild crop relatives can be based on existing DArT platforms developed for the related crop. Genetic diversity in *Aegilops tauschii*, a wild species and a donor of the D genome of wheat, was investigated. Sohail et al. (2012) used 5500 preselected clones from a DArT array developed for wheat, and added 2000 clones obtained de novo from 81 accessions of *A. tauschii*. Almost 70 % markers from the wheat DArT array were polymorphic, while only 34 % of the newly developed *A. tauschii*-specific clones revealed polymorphisms in the diversity collection. A relatively high level of intraspecific genetic diversity was observed. Three groups were observed, generally reflecting their geographic origin and also, at least to some extent, their classification into subspecies. The research allowed identification of accessions that could contribute tolerance to abiotic stresses for wheat breeding. A DArT platform was also developed for einkorn wheat (*Triticum monococcum*) closely related to *Triticum urartu*, a donor of the A genome of the hexaploid wheat. Genetic diversity of 16 *T. monococcum* accessions revealed population structure partially correlating with their genetic and geographic origin (Jing et al. 2009).

Wild *Solanum* species, *Solanum bulbocastanum* and *Solanum commersonii*, close relatives of potatoes and tomatoes, were investigated using the DArT system and revealed a fine microscale genome structural divergence between wild and cultivated species in Solanaceae (Traini et al. 2013).

Applicability of the DArT system is not limited to higher plants. James et al. (2008) developed DArT platforms for genotyping a diploid fern *Asplenium viride* and a haploid moss *Garovaglia elegans*. Sixteen accessions representing each species were investigated for genetic diversity with respect to substrate specificity and geography, respectively. It was shown that intraspecific diversity structure revealed by DArT markers could have been explained by substrate specificity and phylogeographic patterns. The authors indicated possible applications of DArTs in evolutionary investigations, e.g., adaptive radiations, population dynamics, hybridization, introgression, ecological differentiation, and phylogeography.

11.4 Perspectives

The DArT system effectively complements existing technologies in breeding and genomics, especially for crops with limited resources. Diversity Array Technology markers have been developed for a substantial number of plant species. For some species, projects utilizing DArT markers initially aimed at the development of an efficient tool for genetic mapping and the resulting platforms have not yet been used to study genetic diversity. For the list of available species-specific DArT platforms see www.diversityarrays.com.

With respect to genetic diversity investigations, many authors reported that DArT markers provided information on genetic diversity comparable to or exceeding that achievable with other marker systems, often showing greater discriminatory power which likely could have been attributed to a relatively high number of identified polymorphisms, compared to low throughput systems. However, a few exceptions from this general trend should be mentioned. Hurtado et al. (2008) compared performance of 36 SSR markers, with that of ca. 1000 DArT markers and concluded that the former were relatively better at detecting genetic differentiation in cassava germplasm collections. Generally, DArT markers were reported as having relatively high polymorphism information content (PIC) values, however, they were usually slightly less effective compared to SSR markers, e.g., in hop, DArT markers were less polymorphic and had lower PIC than other marker systems (Howard et al. 2011).

In recent years, several novel high-throughput genotyping strategies were developed. They are based on advantages provided by next-generation sequencing (NGS) platforms (reviewed by Davey et al. 2011) and are highly competitive with respect to 'per data point' cost efficiency. Recently, a modification of the DArT system utilizing NGS rather than hybridization to microarrays for polymorphism detection, called DArTseq™, was proposed. It combines the efficient protocol for genomic complexity reduction employed in the conventional DArT system and the power of genotyping-by-sequencing (GBS) approach based on Illumina short read sequencing. As a result, two score tables are produced, comprising DArT and SNP polymorphisms. It proved to be highly efficient in two recent reports on lesquerella (Cruz et al. 2013) and rice (Courtois et al. 2013), resulting in almost 28,000 and almost 17,000 revealed polymorphisms, respectively. In rice, it was shown that the markers covered the genome relatively evenly (Courtois et al. 2013).

On the other hand, a simple assay for site-specific genotyping may be required for only a few sites of special interest, identified as polymorphic using the DArT system. In principle, any DArT clone can be readily sequenced and used to develop a codominant site-specific marker. An example of the DArT marker conversion protocol was recently reported by Macko-Podgórní et al. (2014) who converted one of several DArT markers differentiating wild and cultivated carrots. A general strategy for the development of DArT marker-derived cleaved amplified polymorphic sequence (CAPS) markers involves the following steps: (1) clone sequencing, (2) mapping on the reference sequence and identification of *Pst*I restriction sites flanking the clone, (3) PCR amplification of fragments comprising both restriction sites with pairs of site-specific primers, (4) digestion of PCR products with *Pst*I, and (5) separation by gel electrophoresis (Fig. 11.2). Upon identification of the restriction site comprising the causative polymorphism, the same protocol can be used for routine site-specific genotyping.

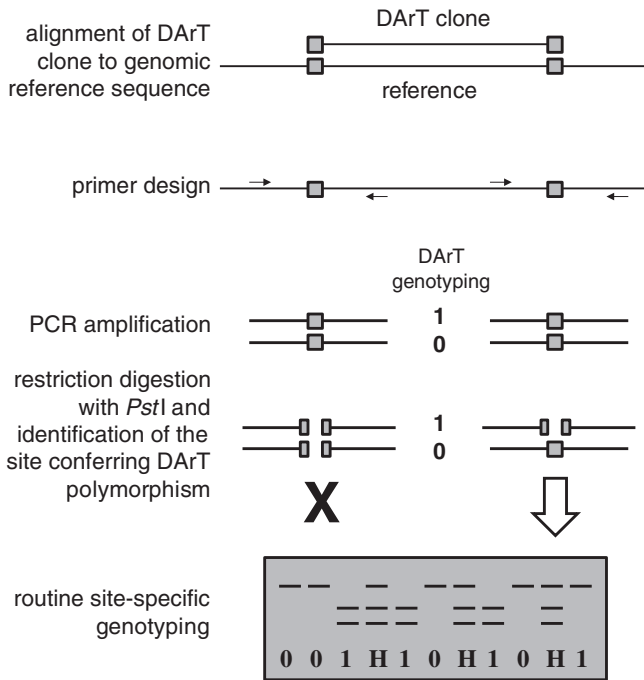


Fig. 11.2 A strategy for conversion of DArT markers into PCR-based codominant site-specific CAPS markers typed by means of gel electrophoresis. *Gray boxes* represent *PstI* restriction sites, *arrows* represent primers, ‘0’, ‘1’, and ‘H’ indicates genotyping scores

11.5 Conclusions

Diversity Arrays Technology was the first high-throughput genotyping platform allowing for parallel detection of hundreds to thousands of polymorphisms in a single assay. It facilitated investigations on genetic diversity in many plant species, representing both major and minor crops, and utilization of genetic resources in breeding programs. Despite the fact that DArT markers are binary (i.e., scored as ‘present’ vs. ‘absent’) and dominant, they can be identified in large numbers, resulting in the high discriminatory power. DArT remains a method of choice, in particular for researchers and breeders working with less-studied crops, e.g., those minor on a global scale, but important for local food security (Varshney et al. 2010). Recent technical advances based on the incorporation of the genotyping-by-sequencing approach into the DArT system (DArTseq™) broaden the possibilities of the technology in the era of NGS.

Acknowledgments This study was supported by the Polish Ministry of Science and Higher Education fund for statutory activities for the University of Agriculture in Krakow.

References

- Akbari M, Wenzl P, Caig V et al (2006) Diversity Arrays Technology (DArT) for high-throughput profiling of the hexaploid wheat genome. *Theor Appl Genet* 113:1409–1420
- Amorim EP, Vilarinhos AD, Cohen KO et al (2009) Genetic diversity of carotenoid-rich bananas evaluated by Diversity Arrays Technology (DArT). *Genet Mol Biol* 32:96–103
- Atienza SG, de la Rosa R, Domínguez-García MC et al (2013) Use of DArT markers as a means of better management of the diversity of olive cultivars. *Food Res Int* 54:2045–2053
- Badea A, Eudes F, Graf RJ et al (2008) Phenotypic and marker-assisted evaluation of spring and winter wheat germplasm for resistance to fusarium head blight. *Euphytica* 164:803–819
- Baird JH, Kopecký D, Lukaszewski AJ et al (2012) Genetic diversity of turf-type tall fescue using Diversity Arrays Technology. *Crop Sci* 52:408–412
- Bolibok-Braĝoszewska H, Heller-Uszyńska K, Wenzl P et al (2009) DArT markers for the rye genome—genetic diversity and mapping. *BMC Genom* 10:578
- Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am J Hum Genet* 32:314–331
- Briñez B, Blair M, Kilian A et al (2012) A whole genome DArT assay to access germplasm collection diversity in common beans. *Mol Breed* 30:181–193
- Courtois B, Audebert A, Dardou A et al (2013) Genome-wide association mapping of root traits in a Japonica rice panel. *PLoS ONE* 8:e78037
- Crossa J, Burgueño J, Dreisigacker S et al (2007) Association analysis of historical bread wheat germplasm using additive genetic covariance of relatives and population structure. *Genetics* 177:1889–1913
- Cruz VMV, Kilian A, Dierig DA (2013) Development of DArT marker platforms and genetic diversity assessment of the U.S. collection of the new oilseed crop *Lesquerella* and related species. *PLoS ONE* 8:e64062
- Davey JW, Hohenlohe PA, Etter PD et al (2011) Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Rev Genet* 12:499–510
- Diatchenko L, Lau YF, Campbell AP et al (1996) Suppression subtractive hybridization: a method for generating differentially regulated or tissue-specific cDNA probes and libraries. *Proc Natl Acad Sci USA* 93:6025–6030
- Domínguez-García MC, Belaj A, de la Rosa R et al (2012) Development of DArT markers in olive (*Olea europaea* L.) and usefulness in variability studies and genome mapping. *Sci Hortic* 136:50–60
- Grzebelus D, Iorizzo M, Senalik D et al (2014) Diversity, genetic mapping, and signatures of domestication in the carrot (*Daucus carota* L.) genome, as revealed by Diversity Arrays Technology (DArT) markers. *Mol Breed* 33:625–637
- Gupta PK, Rustgi S, Mir RR (2008) Array-based high-throughput DNA markers for crop improvement. *Heredity* 101:5–18
- Hang Vu TT, Lawn RJ, Bielig LM et al (2012) Development and initial evaluation of diversity array technology for soybean and mungbean. *Euphytica* 186:741–754
- Heller-Uszyńska K, Uszynski G, Huttner E et al (2011) Diversity Arrays Technology effectively reveals DNA polymorphism in a large and complex genome of sugarcane. *Mol Breed* 28:37–55
- Howard EL, Whittock SP, Jakše J et al (2011) High-throughput genotyping of hop (*Humulus lupulus* L.) utilising Diversity Arrays Technology (DArT). *Theor Appl Genet* 122:1265–1280
- Hurtado P, Olsen KM, Buitrago C et al (2008) Comparison of simple sequence repeat (SSR) and diversity array technology (DArT) markers for assessing genetic diversity in cassava (*Manihot esculenta* Crantz). *Plant Genet Resour Charact Utilization* 6:208–214
- Jaccoud D, Peng K, Feinstein D et al (2001) Diversity arrays: a solid state technology for sequence information independent genotyping. *Nucleic Acids Res* 29:e25
- James KE, Schneider H, Ansell SW et al (2008) Diversity Arrays Technology (DArT) for pan-genomic evolutionary studies of non-model organisms. *PLoS ONE* 3:e1682

- Jing H-C, Bayon C, Kanyuka K et al (2009) DArT markers: diversity analyses, genomes comparison, mapping and integration with SSR markers in *Triticum monococcum*. BMC Genom 10:458
- Kilian A, Huttner E, Wenzl P et al (2005) The fast and the cheap: SNP and DArT-based whole genome profiling for crop improvement. In: Tuberosa R, Phillips RL, Gale M (eds), Proceedings of the international congress “in the wake of the double helix: from the green revolution to the gene revolution”, Bologna, May 2003, Avenue Media, pp 443–461
- Lezar S, Myburg AA, Berger DK et al (2004) Development and assessment of microarray-based DNA fingerprinting in *Eucalyptus grandis*. Theor Appl Genet 109:1329–1336
- Mace ES, Xia L, Jordan DR et al (2008) DArT markers: diversity analyses and mapping in *Sorghum bicolor*. BMC Genom 9:26
- Macko-Podgórní A, Iorizzo M, Smółka K et al (2014) Conversion of a Diversity Arrays Technology marker differentiating wild and cultivated carrots to a co-dominant cleaved amplified polymorphic site marker. Acta Biochim Pol 61:19–22
- Ovesná J, Kučera L, Vaculová K et al (2013) Analysis of the genetic structure of a barley collection using DNA diversity array technology (DArT). Plant Mol Biol Rep 31:280–288
- Powell W, Machray G, Provan J (1996) Polymorphism revealed by simple sequence repeats. Trends Plant Sci 1:215–222
- Rafalski A (2002) Applications of single nucleotide polymorphisms in crop genetics. Curr Opin Plant Biol 5:94–100
- Raman H, Raman R, Nelson MN et al (2012) Diversity Arrays Technology markers: genetic diversity analyses and linkage map construction in rapeseed (*Brassica napus* L.). DNA Res 19:51–65
- Risterucci A-M, Hippolyte I, Perrier X et al (2009) Development and assessment of Diversity Arrays Technology for high-throughput DNA analyses in *Musa*. Theor Appl Genet 119:1093–1103
- Soahil Q, Shehzad T, Kilian A et al (2012) Development of diversity array technology (DArT) markers for assessment of population structure and diversity in *Aegilops tauschii*. Breed Sci 62:38–45
- Tinker NA, Kilian A, Wight CP et al (2009) New DArT markers for oat provide enhanced map coverage and global germplasm characterization. BMC Genom 10:39
- Traini A, Iorizzo M, Mann H et al (2013) Genome microscale heterogeneity among wild potatoes revealed by Diversity Arrays Technology marker sequences. Int J Genom. doi:10.1155/2013/257218
- Varshney RK, Glaszmann J-C, Leung H et al (2010) More genomic resources for less-studied crops. Trends Biotechnol 28:452–460
- Vos P, Hogers R, Bleeker M et al (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23:4407–4414
- Wenzl P, Carling J, Kudrna D et al (2004) Diversity Arrays Technology (DArT) for whole-genome profiling of barley. Proc Natl Acad Sci USA 101:9915–9920
- White J, Law JR, MacKay KJ et al (2008) The genetic diversity of UK, US and Australian cultivars of *Triticum aestivum* measured by DArT markers and considered by genome. Theor Appl Genet 116:439–453
- Williams JGK, Kubelik AR, Livak KJ et al (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res 18:6531–6535
- Wittenberg AHJ, van der Lee T, Cayla C et al (2005) Validation of the high-throughput marker technology DArT using the model plant *Arabidopsis thaliana*. Mol Genet Genom 274:30–39
- Xia L, Peng K, Yang S et al (2005) DArT for high-throughput genotyping of cassava (*Manihot esculenta*) and its wild relatives. Theor Appl Genet 110:1092–1098
- Xie Y, McNally K, Li C-Y et al (2006) A high-throughput genomic tool: diversity array technology complementary for rice genotyping. J Integr Plant Biol 48:1069–1076
- Yang S, Pang W, Ash G et al (2006) Low level of genetic diversity in cultivated *Pigeonpea* compared to its wild relatives is revealed by Diversity Arrays Technology. Theor Appl Genet 113:585–595
- Zalapa JE, Cuevas H, Zhu H et al (2012) Using next-generation sequencing approaches to isolate simple sequence repeats (SSR) loci in the plant sciences. Am J Bot 99:193–208