Theor Appl Genet (2000) 100:633–640

S. Erschadi · G. Haberer · M. Schöniger
R.A. Torres-Ruiz

Estimating genetic diversity of *Arabidopsis thaliana* ecotypes with amplified fragment length polymorphisms (AFLP)

Received: 18 June 1999 / Accepted: 28 July 1999

Abstract The extensive natural variation of *Arabidopsis thaliana* ecotypes is being increasingly exploited as a source of variants of genes which control (agronomically) important traits. We have subjected 19 different *Arabidopsis thaliana* ecotypes to an analysis using the amplified fragment length polymorphism (AFLP) technique in order to estimate their genetic diversity. The genetic diversity was estimated applying the method of Nei and Li (1979) and a modified version of it and using 471 informative polymorphisms. The data obtained revealed that within this small set of ecotypes a group of three ecotypes and a further single ecotype exhibit considerable genetic diversity in comparison to the others. These ecotypes clustered at positions significantly separated from the bulk of the ecotypes in the generated similarity plots. The analysis demonstrated the usefulness of the AFLP method for determinating intraspecies genetic diversity as exemplified with *Arabidopsis thaliana* ecotypes. Results are discussed and compared with data obtained with other methods.

Key words *Arabidopsis thaliana* · Ecotypes · Genetic diversity · AFLP

Introduction

Arabidopsis thaliana as one of the most intensively studied plants (e.g. Anderson and Roberts 1998, for review) has been collected at numerous places all over the world leading to the accumulation of a large number of collected ecotypes [Nottingham Arabidopsis Stock Center (NASC) and Arabidopsis Biological Resource Center (ABRC) Ohio]. Ecotypes are populations that have maintained their identity through isolation or selection in a specific environment (King et al., 1993), but the term "ecotype" is also loosely used for any *Arabidopsis* line collected and preserved through selfing.

Important developmental and physiological traits often exhibit considerable variation in their expression. Therefore, it is not surprising that increasing attention is being paid to the analysis of their quantitative character under aspects of population genetics, ecology and evolution, the more so since many of these traits, such as flowering time, seed size and content, resistance to pathogens etc., are of direct economical importance. The *Arabidopsis* ecotype collection potentially represents a broad source of natural variation embracing a spectrum of variants for any given gene and thus will prove useful in extending the possibilities of this plant as a model system for applied plant molecular biology research. The exploitation of this variability can be followed by different routes. While mapping populations derived from crosses between standard ecotypes like Columbia and other ecotypes are being used to identify and isolate genes responsible for diverse traits (Lister and Dean 1993; Alonso-Blanco et al. 1998), it will certainly also be advantageous to identify the more genetically diverse ecotypes among those of the present collection. Such ecotypes will immediately provide material for establishing highly polymorphic ecotypes and will possibly also exhibit a variation of interesting characters. Moreover, it might be possible in the future to identify polymorphisms which are closely linked to an important trait once a comprehensive overview of genetic diversity data is available.

Genetic distances between *Arabidopsis* ecotypes have been measured by analyses of microsatellites, (preselected) restriction fragment length polymorphism (RFLP) probes or cleaved amplified polymorphic sequence

Communicated by L. Willmitzer

S. Erschadi · G. Haberer · R.A. Torres-Ruiz (✉)
Institut für Genetik, Technische Universität München,
Lichtenbergstr. 4, 85747 Garching, Germany
Fax: +49-89-2891-2892

M. Schöniger
Lehrstuhl für Theoretische Chemie,
Technische Universität München, Lichtenbergstr. 4,
85747 Garching, Germany

S. Erschadi, G. Haberer,
These authors contributed equally to this work

(CAPS) markers and sequences (King et al. 1993; Hanfstingl et al. 1994; Innan et al. 1996, 1997; Hardtke et al. 1996; Ullrich et al. 1997; Bergelson et al. 1998; Purugganan and Suddith 1999). However, due to the markers and methods selected only a limited number of polymorphisms could be analysed. The investigation presented here explored the genetic diversity of *Arabidopsis* ecotypes using the amplified fragment length polymorphism (AFLP) technique and evaluated the use of this technique in such analyses. The analysis has lead to the detection of ecotypes which exhibit significant genetic diversity among 19 different ecotypes.

Materials and methods

Plant material

The source of the *Arabidopsis thaliana* seeds and abbreviations of the ecotypes are listed in Table 1. Seeds were sown on soil in pods (5×5 cm), and plants were grown as described in Torres Ruiz et al. (1996).

Table 1 Analysed *Arabidopsis* ecotypes

Ecotype	Source[a] (order no.)	Geographic origin/location
Ba-1	NASC (N952)	Blackmount/UK
Bi-1	TD	Bickendorf (Cologne)/Germany
Bi-3	TD	Bickendorf (Cologne)/Germany
Bi-5	TD	Bickendorf (Cologne)/Germany
Col-4	NASC/ABRC (N933)	Used for Lister/Dean RI-lines
En-2	NASC (N1138)	Frankfurt/Germany
Est-1	NASC (N1150)	Estland
Hi-0	NASC (N1226)	Hilversum/Netherlands
Kas-1	NASC (N1264)	Kashmir/India
L*er*	NASC/ABRC (NW20)	Used for Lister/Dean RI-lines
L1-0	AS	Llagostera/Spain
Ma-0	AS	Marburg (Lahn)/Germany
Ms-0	AS	Moscow/Russia
Mt-0	TD	Maturba (Cyrenaika)/Lybia
Nd-0	GJ	Niederzenz/Germany
No-0	AS	Nossen=Niederzenz?/Germany
Oy-0	TD	Oystese/Norway
RLD1	NASC (N913)	M. Koornneef/Netherlands
Tsu-0	NASC (N1564)	Tsu/Japan
Ws	GJ	Wassilewskija/Russia

[a] AS, A. Schaeffner (GSF, München); GJ, G. Jürgens (University of Tübingen); TD, T. Debener (BAZ, Ahrensburg); NASC, Nottingham Arabidopsis Stock Center; ABRC, Arabidopsis Biological Resource Center

Isolation of DNA/AFLP reactions

DNA was extracted from pooled plants of each ecotype. We found that the quality of DNA can significantly influence the AFLP banding pattern. Therefore, DNA was isolated from all ecotypes on a large scale according to conventional methods and subjected to a CsCl-gradient purification step. AFLP analysis was essentially performed according to Zabeau and Vos (1993) and Vos et al. (1995) using 500 ng of DNA per ecotype for all AFLP reactions. Briefly, the four-cutter restriction enzyme, *Mse*I, the six-cutter restriction enzyme, *Eco*RI and appropriate adaptors were used for all reactions. AFLP reactions were carried out using a primer specific to the *Eco*RI adaptor with two additional selective nucleotides (E+2) and a primer specific to the *Mse*I adaptor with three additional (in one case two) selective nucleotides (M+3 and M+2, respectively). Ecotypes were genotyped with four different E+2 and nine different M+3/M+2 primers arranged in 15 combinations (see Table 2). E+2 primers were end-labelled with γ-[^{32}P]-ATP (specific activity 110 TBq/mmol. Amersham). Polymerase chain reactions (PCR) were carried out in a Biometra Thermocycler, and amplified fragments were separated on 5% polyacrylamide sequencing gels. Gels were processed by conventional methods, and AFLP patterns were visualized using a Kodak X-ray film. Only clearly detectable bands, which were evaluated individually by two persons by eye, were included in this analysis.

Estimation of genetic similarity

A total of 499 well-separated, polymorphic bands were included into the genetic diversity estimation. Each band was treated as a unit character by scoring its presence (1) or absence (0). From 16 to 60 polymorphic bands per gel were scored depending on the primer combination used (see text). Of the 499 bands 471 were assumed to be informative and included in the similarity estimation (see below). Pairwise genetic similarity was calculated using a method suggested by Nei and Li (1979):

$$GS\ (a, b)=P_{ab}/N_{ab},$$

where P_{ab} is the number of bands present in line a as well as in line b, and N_{ab} is the number of bands present in line a or in line b. In addition, a modified version of Nei and Li's formula was used:

$$GS\ (a, b)=(P_{ab}+A_{ab})/N,$$

where A_{ab} is the number of bands absent in line a as well as in line b, while present in at least 1 of the remaining lines (i.e. ecotypes). N is the total number of bands included into the study (471).

Principal component analysis (PCA) was performed using software from the StatLib-Multivariate Archive (http://lib.stat.cmu.edu/multi/) with the implementation of the GS measures from above. PCA is a standard tool of ordination in numerical taxono-

Table 2 Nomenclature of selective primers and primer combinations used in the AFLP analysis

Primer ↓	→ Selective nucleotides ↓	→ E4 AC	E6 TG	E8 TT	E9 GG
M1	GAA			E8M1 (44/27)	E9MI (29/15)
M3	ACG	E4M3 (25/9)[a]	E6M3 (32/6)	E8M3 (48/14)	
M4	TG				E9M4 (60/19)
M5	CAG	E4M5 (21/15)			E9M5 (18/9)
M6	TCT				E9M6 (23/14)
M7	CTC		E6M7 (39/12)		
M8	TTG		E6M8 (23/14)	E8M8 (55/23)	E9M8 (16/8)
M9	TAG		E6M9 (31/9)		
M11	CTA	E4M11 (35/18)			

[a] Numbers in parentheses indicate number of polymorphic/non-polymorphic bands obtained

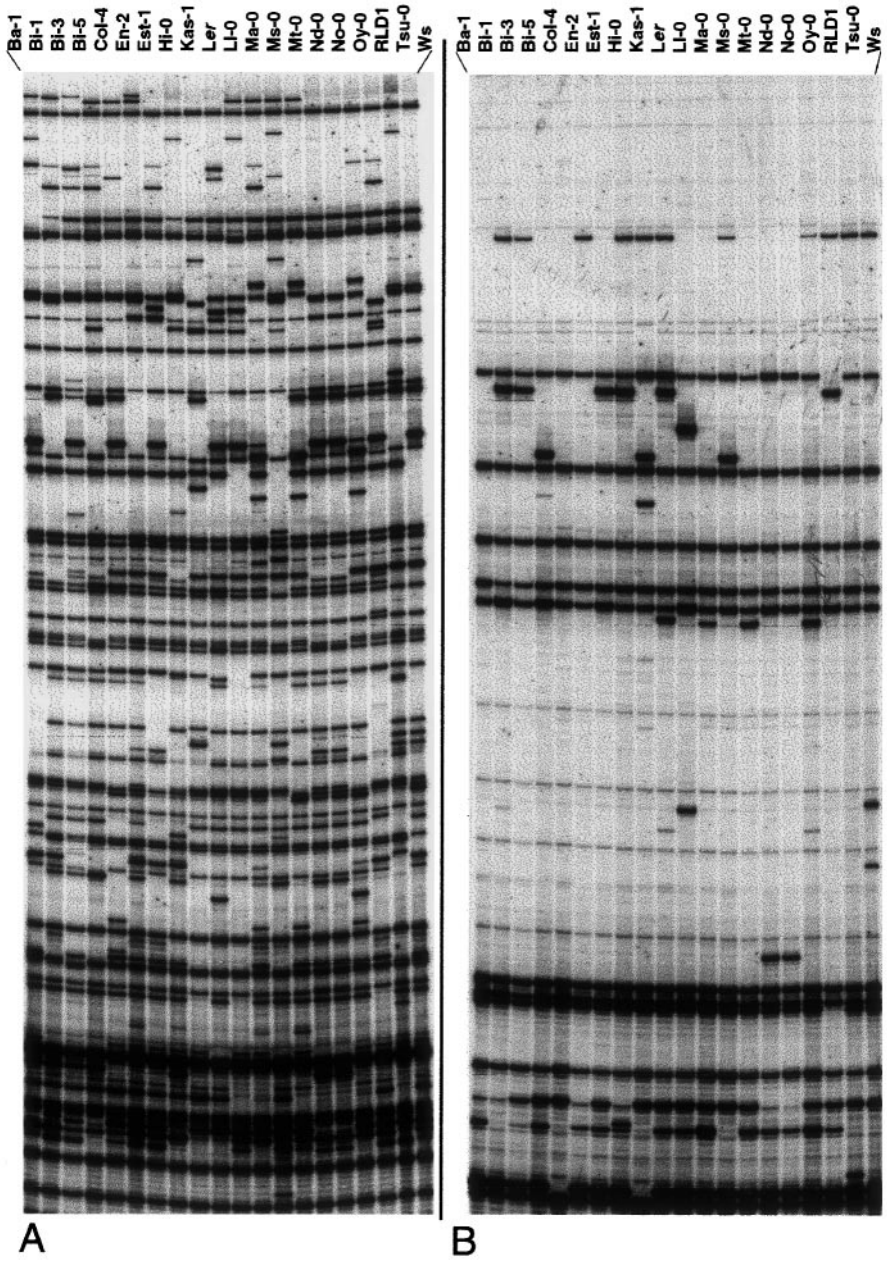

Fig. 1A, B Representative AFLP fingerprints of different *A. thaliana* ecotypes generated with primer combination E9M4 (**A**) and E9M8 (**B**), respectively

my (Sneath and Socal 1973). It allows the relationships between different objects (ecotypes) to be visualized economically in a two-dimensional plot. The underlying solution of the eigenvalue problem provides a transformation of a hyperspace such that only the first (two) dimensions are sufficient to represent the majority of variance present in the data. A bootstrap analysis was performed to estimate the variance of genetic similarity values. Two hundred samples of size 50, 100, 150, …, 450 bands were randomly drawn from the 471 polymorphic bands assumed to be informative, and the GS values were calculated for all 171 pairs of ecotypes using Nei-Li's formula and the modified version of it. For each sample the coefficient of variation (CV) was calculated: CV=standard deviation/sample mean.

Subsequently, the CVs were averaged over all 200 replicates.

Results

Evaluation of AFLP fingerprints

Twenty *Arabidopsis* ecotypes (Table 1) were analysed with 15 AFLP primer combinations. No difference was detected between Nd-0 (Niederzenz) and No-0 (Nossen). It is possible that No-0 was confounded in the process of amplification in the laboratory or long before reception. Note that No-0 has been designated as Niederzenz (Table 1 in Frank et al. 1998). Thus, No-0 was not taken into consideration in subsequent experiments. All primer combinations except 1 (combination E9M4) carried five selective nucleotides. With respect to a particular primer combination all ecotypes were analysed on the same gel,

such that they could be unambiguously examined with respect to presence or absence of any given band. We counted all clearly recognizable polymorphic bands (a band is polymorphic when it is missing in at least 1 ecotype). The minimum number of total bands counted with 1 particular primer combination was 24 and the maximum number 79 (Fig. 1). In total, we detected 499 polymorphic positions, with an average of 33 per combination (see Table 2 for detailed data on polymorphisms detected). The total number of positions where all ecotypes produced a band was 212. We separately counted polymorphisms for the ecotype pair L*er* and Col-4 since an ample AFLP map has recently been constructed using the same core primers and ecotypes, which allows comparison of the data (Alonso-Blanco et al. 1998; see also NASC/ABRC seed list and http://nasc.nott.ac.uk/contents/ler.html). From the 711 positions analysed in this work using 15 different primer combinations 144 were polymorphic between L*er* and Col-4, 65 were specific for L*er* and 79 were specific for Col-4; 567 bands were shared between both ecotypes.

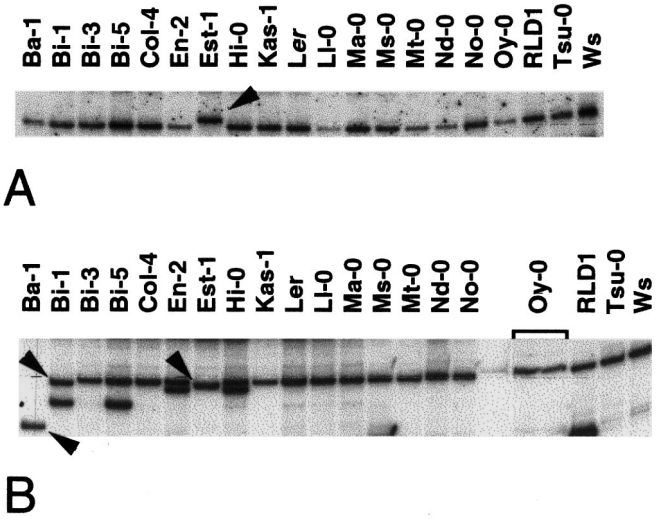

Fig.2A, B Types of complementary AFLP segregation patterns; type I in **A** and type II in **B** (for details see text)

Estimation of frequency of allelic bands in AFLP gels

In contrast to other marker techniques, AFLP analysis allows the simultaneous analysis of 50 or more bands (depending on the number of selective nucleotides used). Consequently it is not possible to discriminate allelic bands amplified from the same locus, i.e. AFLPs are mapped as alleles while RFLPs are mapped as loci, which may result in an overestimation of dissimilarity. Conversely, co-migrating bands (bands at the same gel position) might represent non-homologous loci between the ecotypes leading to an underestimation of dissimilarity.

The latter problem seems to be less serious since work with potato and *Arabidopsis* ecotypes has shown that co-migrating bands are generally allelic (Rouppe van der Voort et al. 1997; Alonso-Blanco et al. 1998). With respect to the former problem Alonso-Blanco et al. (1998) have suggested that 14–20% of all polymorphisms might be allelic between two different *Arabidopsis* ecotypes. Most importantly, the predominant number of AFLP markers naturally segregate in a complementary fashion, and most of these differ in their gel position by a maximum of five bases in *Arabidopsis* (90%, Alonso-Blanco et al. 1998) or one base in potato (Rouppe van der Voort et al. 1997).

Following these criteria the gels were checked for this kind of complementary segregation of AFLP patterns generated by the same primer combination. Two types of patterns, which satisfy the forementioned criteria, were identified (Fig. 2). Type 1 is fully complementary (Fig. 2A). Most ecotypes were non-polymorphic for a given band, which led to a nearly complete line of bands across the gel at a certain position, except for a single or few gaps for those ecotypes which were polymorphic. The gap(s) can be ”filled” with a band(s) only one or a few bases above or beneath the first band and which is/are unique for those ecotypes that missed the first band. Type 2 is very similar to type 1 except that more complex (complementary) patterns were assumed; for instance 1 or some ecotypes were allowed to miss a complementary band completely, while other ecotypes had a complementary band at different nearby positions (Fig. 2B). We found four type-1 and 19 type-2 patterns. This reduced the number of informative polymorphisms to 471 polymorphic bands.

Estimation of genetic diversity of *Arabidopsis* ecotypes

We applied the method of Nei and Li (1979) and a modified version thereof to estimate the genetic similarity/diversity between the ecotypes (see Materials and methods) based on the 471 remaining polymorphic bands. Nineteen ecotypes were included in the analysis, resulting in 171 pair-wise comparisons. The original data and similarity matrices are not shown due to space limitations but can be obtained on request.

The similarity values (SVs) for the 171 ecotype pairs ranged from 0.27 to 0.80 (Nei-Li) and 0.63 to 0.92 (modified version), respectively. When the method of Nei and Li (1979) was applied the first and second principal components of the genetic similarity matrix accounted for 47.2% and 6.0% of the variance, respectively (Fig. 3). In this plot the majority of ecotypes do not show extended dissimilarity; the exceptions are Ma-0, Mt-0, Oy-0 which constitute a small cluster (designated MMO group in the following), and Kas-1, which is an isolated ecotype (Fig. 3). It should be noted that the similarity information in the plot according to this method only accounted for about 53% of the variance and thus excluded a considerable part of the total information. However, in the modified version of Nei and Li applied (see Materials and methods) the first and second

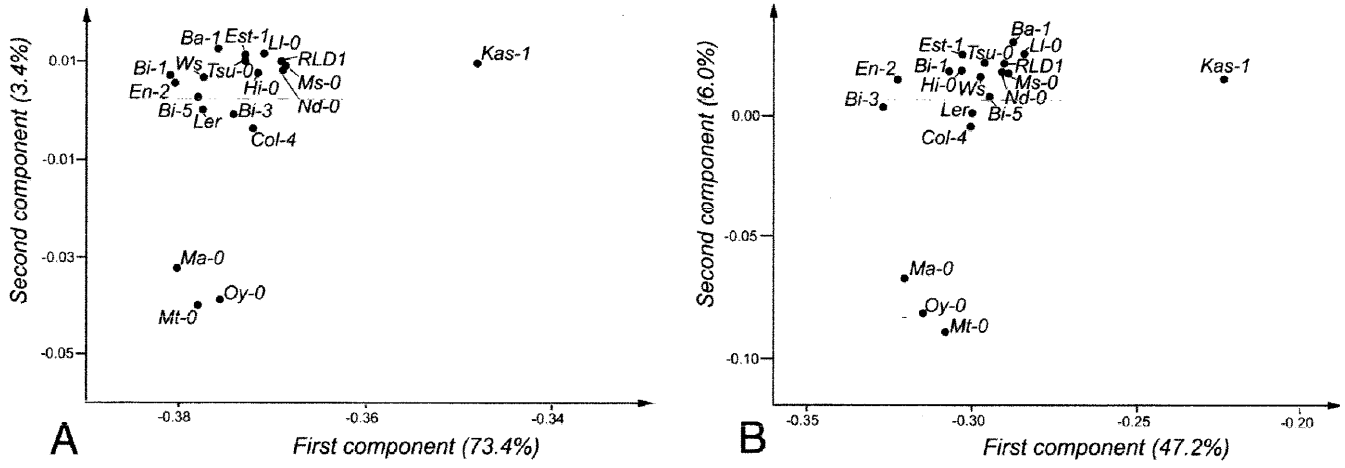

Fig. 3 Plots of first and second principal components of genetic similarity/diversity among 19 *A. thaliana* ecotypes. Plot generated with the modified (**A**) and Nei-Li (**B**) methods, respectively. The *numbers in parentheses* represent the percentage of variation accounted for by each principal component

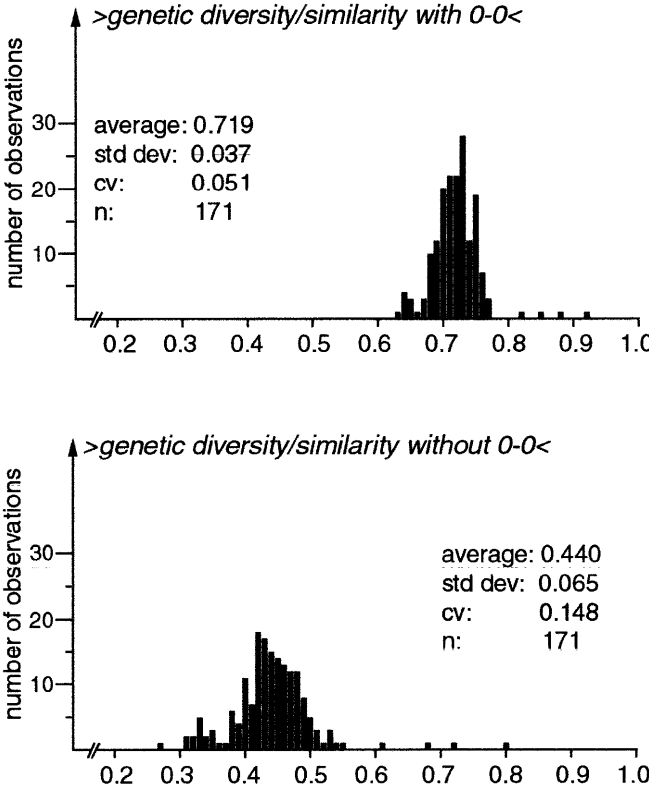

Fig. 4 Distribution of genetic similarity/diversity values among *A. thaliana* ecotypes according to the modified (*top*) and Nei-Li (*bottom*) methods respectively

principal component of the genetic similarity matrix accounted for 73.4% and 3.4% of the variance, respectively. Again, most of the ecotypes are found in one cluster (Fig. 3). The MMO group and Kas-1 respectively are separated from all others. This procedure leads to a stronger compression of the pair-wise similarity values around a mean in comparison to the original Nei/Li procedure (see Fig. 4). Other versions of the genetic similarity estimation (not shown, available upon request) were applied and did also not essentially alter this picture.

The separation of the mentioned ecotype groups also became apparent when particular diversity/similarity values of the Nei-Li or the modified matrix were considered. According to the first method the most distantly related ecotypes are Kas-1 and Ba-1 (SV=0.27, Fig. 4). The second method leads to Kas-1 and Bi-3 as the most distant ecotypes (SVs=0.63, Fig. 4). Of the 17 values below SV=0.375 of the Nei-Li matrix 16 are occupied by pairs where Kas-1 is one member. The same applies for 12 out of 18 values below SV=0.68 in the modified matrix. Interestingly with both methods the highest similarity values were found within the MMO group and ranged from 0.68 to 0.80 (Nei-Li, Fig. 4) and 0.85 to 0.92 (modified, Fig. 4), respectively. Also with both methods, the ecotypes with the highest similarity next to the MMO group were Est-1 and RLD1 (0.61 and 0.82, respectively). These values became more interesting when the Bi ecotypes were taken into consideration. The Bi ecotypes originate from sub-populations of the same location (city district Bickendorf/Cologne/Germany). They are clearly distinguishable from each other but fall into the large ecotype cluster in the similarity plots (Fig. 3). Their values of pair-wise comparisons were moderate with both methods: Bi-1 versus Bi-3 (0.54 and 0.76, respectively), Bi-1 versus Bi-5 (0.46 and 0.75, respectively) and Bi-3 versus Bi-5 (0.53 and 0.76, respectively). The bootstrap analysis indicated that 471 polymorphic bands is an appropriate number for estimating genetic similarity among the ecotypes since a CV of less than 6% is already achieved with the evaluation of 200 bands (application of the modified version, results not shown). In a similar study with RFLPs only 55 bands were estimated to be sufficient for similarity analysis with CVs of 10% (King et al. 1993).

Discussion

This work explored the genetic diversity between *Arabidopsis thaliana* ecotypes by evaluating AFLP patterns. Because of the method applied, the nuclear DNA was isolated together with the chloroplast (cp) and mitochondrial (mt) genomes. However, it is likely that the polymorphisms generated predominantly originated from the nuclear genome, except for those which fall into a limited number of rearrangement breakpoints in mitochondrial genomes (Palmer and Herbon 1988). The evolutionary divergence rate of plant organellar DNA components has been found to be significantly lower than the nuclear rate (Wolfe et al. 1987; Palmer and Herbon 1988). Plant mtDNA evolves three to four times slower than cpDNA, which in turn evolves at least twice (the large inverted cpDNA repeat six times) as slowly as nuclear DNA (Wolfe et al. 1987; Palmer and Hebron 1988). Bergelson et al. (1998) did not find any polymorphism in the mitochondrial Nad5 locus between 115 field-collected lines and 7 additional ecotypes. Since the *Arabidopsis* nuclear genome harbours only a low amount of repetitive elements (Leutwiler et al. 1984; Pruitt and Meyerowitz 1986) it is probable that many of the strong non-polymorphic signals/bands in the AFLP gels belong to organellar genomes.

AFLP analysis as a method for analysing genetic diversity: an assessment

The AFLP method is robust and reproducible as revealed by comparison with the mapping work of Alonso-Blanco et al. (1998). The number of non-polymorphic bands between L*er* and Col-4 was virtually identical, i.e. 35.8 (average numbers, excluding E9M4 in our work) versus 33.6. Polymorphisms were significantly underrepresented in our study (9.4 versus 14.7, average numbers excluding E9M4 in our work). This is probably caused by the use of [^{32}P] in this work, while Alonso-Blanco et al. (1998) used [^{33}P], which is known to produce sharper and clearer bands. In the present work fuzzy polymorphic bands were deliberately not counted. Such bands would be sharp and their existence unambiguously controlled by mapping in the work of Alonso-Blanco et al. (1998).

Two problems arise when genetic diversity between closely related organisms is measured by the AFLP method. Firstly, diversity might be underestimated by assuming non-polymorphic bands to be allelic. It seems that this problem can at least partly be neglected. In a mapping study including 3 *Arabidopsis thaliana* ecotypes it has been shown that bands of similar molecular size between two ecotypes but polymorphic with the third represent the same locus by virtue of their estimated map position (Alonso-Blanco et al. 1998). Rouppe van der Voort et al. (1997) have shown directly by sequence analysis of potato AFLP bands that those generated with the same primer combination at the same gel position are allelic between different accessions. The same authors calculated a frequency of 0.03 for coincident co-migration arising by chance. Secondly, genetic diversity might be overestimated by considering some polymorphic bands to be non-allelic. The work with potato accessions has shown that polymorphic bands can be allelic when they occur at nearby positions, i.e. positions differing by one base (Rouppe van der Voort et al., 1997). Alonso-Blanco et al. (1998) have shown for the L*er*/Col-4 and L*er*/Cvi pairs that 14% and 20% of the bands, respectively, might be allelic by map position. Most importantly, these AFLP markers naturally segregate in a complementary fashion and most of them (90%) differ in their gel position by a maximum of five bases (Alonso-Blanco et al. 1998). We have used this information to find such complementary AFLP patterns between ecotypes and found about 5% to be possible candidates. This significantly smaller percentage compared to that found by Alonso-Blanco et al. (1998) might result from viewing more complex complementary patterns, which arise when many ecotypes are used. It should also be noted that bands considered to be allelic, by map position might well represent non-allelic closely linked loci. Taken together, both problems counterbalance each other with the overestimation of diversity being perhaps more prominent (note additional counterbalance to this effect by the modified Nei-Li method, see below).

Detection of genetically diverse *Arabidopsis* ecotypes

Diversity/similarity plots were obtained by applying the method of Nei and Li (1979) and a modification of it. The latter takes into account additional similarity between ecotypes which lack bands at polymorphic positions (see Materials and methods). This counteracts the possible overestimation of diversity discussed above in the following way. It disregards the fact that the missing bands might also show additional polymorphisms, an effect which tends to underestimate diversity. It should also be noted that this procedure makes sense because only 15 out of the possible 1024 primer combinations with five selective nucleotides were tested, otherwise the missing bands of these ecotypes surely would have been captured with another primer combination in another gel (or even the same, see above). In fact, both methods produce essentially the same result (Figs. 3, 4). The data point to ecotypes with significant genetic divergence. Three clusters are separated from each other. The largest cluster embraces 15 ecotypes, including standard lines, i.e. Col-4, L*er*, Nd-0 and Ws. A second small cluster consists of the MMO group (Ma-0, Mt-O, Oy-0), while Kas-1 represents an isolated ecotype. Additionally, the pair-wise comparisons of ecotypes (not shown) point to considerably polymorphic pairs.

Comparison with data obtained with other molecular markers

Are these results reasonable in context of other analyses? Other studies have used selected RFLP probes, CAPS markers, (gene) sequences and microsatellites to compare *Arabidopsis* ecotypes (King et al. 1993; Hanfstingl et al. 1994; Hardtke et al. 1996; Innan et al. 1996, 1997; Ullrich et al. 1997; Bergelson et al. 1998; Purugganan and Suddith 1999). Due to the different material and, in some cases, low numbers of polymorphisms evaluated only basic features can be compared at present. Essentially, in all analyses most ecotypes are not significantly separated from each other. In accordance with all other studies the AFLP analysis did not reveal a link between geographical location and genetic diversity although it did reveal a substantial difference between populations collected from one location (Bi ecotypes). A similar observation was made by Bergelson et al. (1998) and indicates a low interpopulation migration rate. The genetic difference between Col-4 and L*er* does not point to a general Col/L*er* dichotomy as found in the Adh gene (Hanfstingl et al. 1994; Innan et al. 1996) and which has been interpreted as a special case of balanced selection of two Adh haplotypes (Hanfstingl et al. 1994). In cases where Kas-1 and other ecotypes used in this study have been included in other studies, the former ecotype is clearly separated from the latter (Hardtke et al. 1996; Innan et al. 1997; Purugganan and Suddith 1999). Hardtke et al. (1996) used many of the ecotypes analysed in this study and found the MMO group (together with Est-0 and Col-3) to be well separated from all other ecotypes when nuclear sequences were analysed but not when two chloroplast sequences were included. In both studies Est-1 and RLD were closely related (nuclear and chloroplast genes included). The ecotype Tsu-0 is particularly interesting because, together with 4 other Japanese ecotypes, it does not cluster within the otherwise distinct Japanese ecotype group found by evaluating microsatellite variation (Innan et al. 1997). Its elevated similarity with non-Japanese ecotypes has been found in all studies. One discrepancy between the studies concerns a more or less marked separation of Col-0 or Nd-0 from other ecotypes, for instance L*er*. Undoubtedly, the history of L*er* and Col suggests that both ecotypes should indeed be different (http://nasc.nott.ac.uk/contents/ler.html). However, this is unlikely to account for large differences, particularly since both originated from the same population. One source of difference might result from using different lines/accessions from 1 ecotype, e.g. No-0 (UK) versus No-0 (USA) (Frank et al. 1998) or Col-4 versus Col-0. We suspect that additional (biased) dissimilarity was obtained by using probes/markers which had been deliberately selected for being polymorphic between the standard *Arabidopsis thaliana* strains: Col, L*er* and Nd. This is the case for the RFLP probes isolated by Chang et al. (1988) and the CAPS markers from Konieczny and Ausubel (1993). One source of bias can also origin from a inappropriate number of informative sites being evaluated. The power of AFLP can be estimated by considering the number of polymorphisms generated, which is the highest of all studies so far. Logically, this elevates the probability of finding truly divergent ecotypes. Interestingly, Mt-0 and Oy-0 are among those ecotypes with the lowest ability to form auxin-induced adventitious roots (King and Stimart 1998), but it remains to be shown how far elevated polymorphism frequencies correlate with bio-diversity. Taken together we conclude that AFLP is a strong and valuable tool with which to explore and analyse genetic diversity within *Arabidopsis thaliana* as within other species. The present study has detected a few cases with significant genetic diversity within a small subset of all known ecotypes.

Acknowledgements We thank M. Frey and T. Golds for critical reading of the manuscript and G. Jürgens, T. Debener and T. Schäffner as well as the *Arabidopsis* Stockcenters in Ohio and Nottingham for seeds. We thank especially A. Gierl for consistent and generous support of our work. This work was partially supported by the Deutsche Forschungsgemeinschaft (R.A.T.-R.) and a studentship from the Technical University (G.H). All experiments performed comply with the German laws concerning ”Biologische Sicherheit”.

Note added in proof We would like to point to the work of our colleagues Breyne et al. published in Mol Gen Genet 261:627–634 while this manuscript was under review.

References

Alonso-Blanco C, Peeters AJM, Koornneef M, Lister C, Dean C, van den Bosch N, Pot J, Kuiper MTR (1998) Development of an AFLP based linkage map of Ler, Col and Cvi *Arabidopsis thaliana* ecotypes and construction of a Ler/Cvi recombinant inbred line population. Plant J 14:259–271

Anderson M, Roberts JA (1998) Annual plant reviews vol.1: *Arabidopsis*. Sheffield Academic Press, Sheffield, UK

Bergelson J, Stahl E, Dudek S, Kreitman M (1998) Genetic variation within and among populations of *Arabidopsis thaliana*. Genetics 148:1311–1323

Chang C, Bowman JL, DeJohn AW, Lander ES, Meyerowitz EM (1988) Restriction fragment length polymorphism linkage map for *Arabidopsis thaliana*. Proc Natl Acad Sci USA 85: 6856–6860

Frank MJ, Preuss D, Mack A, Kuhlmann TC, Crawford NM (1998) The *Arabidopsis* transposable element TAG1 is widely distributed among *Arabidopsis* ecotypes. Mol Gen Genet 257:478–484

Hanfstingl U, Berry A, Kellogg EA, Costa JT III, Rüdiger W, Ausubel FM (1994) Haplotypic divergence coupled with lack of diversity at the *Arabidopsis thaliana* alcohol dehydrogenase locus: roles for both balancing and directional selection? Genetics 138:811–828

Hardtke CS, Müller J, Berleth T (1996) Genetic similarity among *Arabidopsis thaliana* ecotypes estimated by DNA sequence comparison. Plant Mol Biol 32:915–922

Innan H, Tajima F, Terauchi R, Miyashita NT (1996) Intragenic recombination in the Adh locus of the wild plant *Arabidopsis thaliana*. Genetics 143:1761–1770

Innan H, Terauchi R, Miyashita NT (1997) Microsatellite polymorphism in natural populations of the wild plant *Arabidopsis thaliana*. Genetics 146:1441–1452

King G, Nienhuis J, Hussey C (1993) Genetic similarity among ecotypes of *Arabidopsis thaliana* estimated by analysis of restriction fragment length polymorphisms. Theor Appl Genet 86:1028–1032

King JJ, Stimart DP (1998) Genetic analysis of variation for auxin-induced adventitious root formation among eighteen ecotypes of *Arabidopsis thaliana* L. Heynh.. J Hered 89:481–487

Konieczny A, Ausubel FM (1993) A procedure for mapping *Arabidopsis* mutations using co-dominant ecotype-specific PCR-based markers. Plant J 4:403–410

Leutwiler LS, Hough-Evans BR, Meyerowitz EM (1984) The DNA of *Arabidopsis thaliana*. Mol Gen Genet 194:15–23

Lister C, Dean M (1993) Recombinant inbred lines for mapping RFLP and phenotypic markers in *Arabidopsis thaliana*. Plant J 4:745–750

Nei M, Li W-H (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc Natl Acad Sci USA 76:5269–5273

Palmer JD, Herbon LA (1988) Plant mitochondrial DNA evolves rapidly in structure, but slowly in sequence. J Mol Evol 28: 87–97

Pruitt RE, Meyerowitz EM (1986) Characterization of the genome of *Arabidopsis thaliana*. J Mol Biol 187:169–183

Purugganan MD, Suddith JI (1999) Molecular population genetics of floral homeotic loci: departures from equilibrium-neutral model at the *APETALA3* and *PISTILLATA* genes of *Arabidopsis thaliana*. Genetics 151:839–848

Rouppe van der Voort JNAM, van Zandvoort P, van Eck HJ, Folkertsma RT, Hutten RCB, Draaistra J, Gommers FJ, Jacobsen E, Helder J, Bakker J (1997) Use of allele specificity of co-migrating AFLP markers to align genetic maps from different potato genotypes. Mol Gen Genet 255:438–447

Sneath PHA, Sokal RR (1973) Numerical taxonomy. The principles and practice of numerical classification. WH Freeman and Co, San Francisco

Torres Ruiz RA, Lohner A, Jürgens G (1996) The *GURKE* gene is required for normal organization of the apical region in the *Arabidopsis* embryo. Plant J 10:1005–1016

Ullrich H, Lättig K, Brennicke A, Knoop V (1997) Mitochondrial DNA variations and nuclear RFLPs reflect different genetic similarities among 23 *Arabidopsis thaliana* ecotypes. Plant Mol Biol 33:37–45

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23:4407–4414

Wolfe KH, Li W-H, Sharp PM (1987) Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast and nuclear DNAs. Proc Natl Acad Sci USA 84:9054–9058

Zabeau M, Vos P (1993) European Patent Application publication number EP 0534858