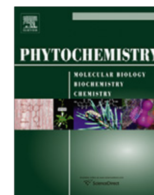
Phytochemistry 72 (2011) 1243–1250

Contents lists available at ScienceDirect

Phytochemistry

journal homepage: www.elsevier.com/locate/phytochem



The use of 2D-electrophoresis and *de novo* sequencing to characterize inter- and intra-cultivar protein polymorphisms in an allopolyploid crop

Sebastien Christian Carpentier [a,*], Bart Panis [a], Jenny Renaut [b], Bart Samyn [c], Annelies Vertommen [a], Anne-Catherine Vanhove [a], Rony Swennen [a], Kjell Sergeant [b]

[a] K.U.Leuven, Department of Biosystems, Division of Crop Biotechnics, Leuven, Belgium
[b] CRP Gabriel Lippmann, Department of Environment and Agrobiotechnologies, Belvaux, Luxembourg
[c] Ghent University, Department of Biochemistry and Microbiology, Gent, Belgium

ARTICLE INFO

Article history:
Available online 23 November 2010

Keywords:
2D-DIGE
Allopolyploid crops
de novo sequencing
Protein polymorphism
Plant proteomics

ABSTRACT

Polyploidy and allopolyploidy have played an important role in the evolution of many plants and crops. Several techniques exist to characterize allopolyploid varieties. Analyzing the consequences of genomic reorganization at the gDNA level is a prerequisite but a better insight into the consequences for the phenotype is also primordial. As such, protein polymorphism analysis is important in understanding plant and crop biodiversity and is a driving force behind crop improvement. Our strategy to analyze protein isoforms and to detect possible gene silencing or deletion in bananas was based on protein analysis. Bananas are a good representative of a complex allopolyploid and important crop. We combined two-dimensional electrophoresis (2DE) and 2D DIGE with *de novo* MS/MS sequence determination to characterize a range of triploid varieties. Via Principal Component Analysis (PCA) and hierarchical clustering we were able to blindly classify the different varieties according to their presumed genome constitution. We report for the first time the application of an automated approach for the derivatization of peptides for facilitated MS/MS *de novo* sequence determination. We conclude that the proteome does not always correspond to the presumed genome formulae and that proteomics is a powerful tool to characterize varieties. The observations at the protein level provide good indications for a more complex genome structure and genomic rearrangement in some banana varieties.

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

Polyploidy and allopolyploidy have played an important role in the evolution of plants including important crops (Soltis and Soltis, 1999). While the presumed advantage of allopolyploidy is an increase in genetic diversity, its occurrence considerably complicates the analysis of allopolyploid crops. Genome reorganization in allopolyploids (both intra- and inter-genomic reorganization of the different genomes) can take place already within the F1 generation (Soltis and Soltis, 1999). Allopolyploid varieties can be characterized via several techniques. DNA technology offers a number of options to determine the genomic content: e.g. (rDNA) (Boonruangrod et al., 2009; Rauscher et al., 2004), Genomic In Situ Hybridization (GISH) (Schwarzacher et al., 1989), cytoplasmic markers (Boonruangrod et al., 2008; Carreel et al., 2002; Corriveau and Coleman, 1988). Attractive approaches to detect large numbers of genome-specific Single Nucleotide Polymorphism (SNP) markers are DArT (Diversity Array Technology) (Wenzl et al., 2004) and (eco)TILLING (Targeting Induced Local Lesions IN Genomes) (Colbert et al., 2001; Gilchrist et al., 2006; Till et al., 2010). The term SNP is used to indicate polymorphisms of a certain genomic locus being a single base insertion, a deletion or a substitution. The biological significance of a SNP differs depending upon its location within the genome. Mutation is the driving force of evolution. If a SNP mutation takes place in the protein encoding region of a gene, then it might cause: (i) a shift or a stop in the open reading frame, (ii) a shift in one amino acid (called Single Amino Acid Polymorphism (SAAP)) and thus a potential change in protein function/activity or (iii) no effect since a SNP substitution can be neutral thanks to the degeneracy of the genetic code (61 codons for 20 amino acids) and hence no change in protein composition takes place. Though advances in high-throughput and computational technologies have resulted nowadays in the sequencing of the genome of hundreds of organisms and though it is indeed extremely useful and necessary to have reference genomes, we also need a technique that is able to analyze existing and new hybrid varieties and to understand their potential. Analyzing the consequences of genomic reorganization at gDNA level is a prerequisite but a good insight at phenotype level is also primordial. As such, protein

* Corresponding author. Address: K.U.Leuven, Faculty of Bioscience Engineering, Division of Crop Biotechnics, Kasteelpark Arenberg 13, 3001 Leuven, Belgium. Tel.: +32 (0)16 321633/421; fax: +32 (0)16 321993.
E-mail address: sebastien.carpentier@biw.kuleuven.be (S.C. Carpentier).

0031-9422/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.phytochem.2010.10.016

polymorphism analysis is important in understanding plant and crop biodiversity and is a driving force behind crop improvement.

Methods for identifying protein polymorphisms are generally based on analyzing the sequence of individual interesting proteins (Wada, 1996). To realize a blind high throughput screening of protein polymorphisms, Liu and Regnier developed an experimental approach based on a heavy isotope coding method (Liu and Regnier, 2002). The method exploits differential derivatization of amine and carboxyl groups. After differentially double labelling samples, all the peptides having the same sequence in the control and experimental samples appear as a doublet in mass spectra. Uniquely different peptides will appear as a singlet. Though this method is efficient in detecting polymorphisms at the peptide level, this method has the disadvantage to lose the connectivity between peptides derived from the same protein, which complicates protein identification tremendously in many unsequenced plants and crops. Proteome research on non-model plants and crops is absolutely feasible and successful but is currently restricted to protein-based gel approaches (Carpentier et al., 2008). The cause of the gel-based restriction is the fact that a confident identification of the proteins can only be achieved through cross-species searching. In the cross-species approach, proteins can only be confidently identified by comparing several of their peptides to the *in silico* predicted peptides of orthologous proteins of characterized species. However, if not enough peptides are conserved this strategy fails. Sequence reconstruction of an unknown peptide based on the acquired mass data is then the only option. One way to facilitate *de novo* sequence determination is by adding a permanent negative charge to the N-terminus of tryptic peptides (*in casu* a sulfonic acid group) (Samyn et al., 2007; Sergeant et al., 2005). By this the positive charge is counterbalanced. As the most basic residue at the C-terminal part is already protonated, excess protons, so called ionizing protons, will be more or less free to randomly ionize backbone amide groups favouring charge directed fragmentation of the weakened peptide bonds. Fragments that contain the negatively charged N terminus (b-ions) are not detected and only C-terminal y-ions will be visible which facilitates *de novo* sequence determination. By simple calculation of mass differences between consecutive peaks the sequence of the peptide can be deduced. For an overview of proteomics strategies for non-model plants and crops (see Carpentier et al., 2008; Vertommen et al., 2010).

Bananas and plantains (*Musa* spp., collectively called bananas) constitute a group of important crops characterized by a complex background of polyploidy and allo(poly)ploidy. Bananas provide a staple food for many millions of people living in the humid tropics. The cultivated banana is a sterile, parthenocarpic plant selected by early farmers in South-East Asia, and thereafter maintained by vegetative propagation (Heslop-Harrison and Schwarzacher, 2007). Most cultivated banana varieties are triploid and originated from intra- and inter-specific hybridizations between seed-bearing subspecies of *Musa acuminata* (A genome donor) and *Musa balbisiana* (B genome donor) (Simmonds and Sheppard, 1955). If no rearrangement takes place, a hybrid variety originating from *M. acuminata* and *M. balbisiana* should produce both A- and B-isoforms of the same protein proportional to the number of A and B-chromosomes. However, it has been shown in many other allopolyploid species that, as a consequence of genome restructuring, gene-level changes and gene silencing and deletion occurs (Soltis and Soltis, 1999). In some polyploids, chromosomal reorganization and gene silencing ('diploidization') are so extensive that the genome is no longer structured as an allopolyploid. Maize is such an example of a partially 'diploidized' polyploidy (Gaut and Doebley, 1997). *Acuminata* and the *balbisiana* species diverged form a common ancestor and both started to evolve. Nowadays, the *acuminata* and *balbisiana* varieties are strikingly different in phenotype. Proteins are one of the main determinants of the phenotype. Evolution resulted in divergence of proteins that became specific for *acuminata* (A isoforms) and proteins that became specific for *balbisiana* (B isoforms). Bananas are classified on the basis of their genome constitution. The most common types are: AA, BB, AB (diploids) and AAA, AAB, ABB and BBB (triploids). Simmonds attempted to classify the triploids based on crosses with the parental species and phenotypical characteristics (Simmonds, 1966).

We (at K.U.Leuven) host the *Musa* International Germplasm Collection (>1200 accessions) as an *in vitro* and cryopreserved collection (Panis et al., 1996; Van den houwe et al., 1995). Our strategy to characterize this biodiversity and to detect protein isoforms and gene silencing (deletion) events as a result of hybridization and rearrangements was to start from a protein based approach. We combined two-dimensional electrophoresis (2DE) and 2D DIGE with the derivatization of peptides for easy *de novo* MS/MS sequence determination. To characterize a range of triploid varieties and to detect typical A and B isoforms and possible deletions or silencing, a 2DE analysis has been performed. We started by analyzing different triploid varieties classified as AAA, AAB, ABB and BBB. Via Principal Component Analysis (PCA) and hierarchical clustering we were able to blindly classify the different varieties according to their presumed genome constitution. To obtain an insight into the origin of the different protein polymorphisms, we also characterized the diploid wildtype varieties *acuminata* (AA) and *balbisiana* (BB). We report for the first time the application of an automated approach for the derivatization of peptides for MS/MS *de novo* identification. We have already proven that this method (Sergeant et al., 2005) is powerful to unravel protein isoforms in banana (Samyn et al., 2007) and have applied it here in a more high throughput way. We could conclude that the proteome does not necessarily correspond to the presumed genome formulae. The observations at the protein level provide additional indications for a more complex genome structure and genomic rearrangement in some banana varieties as we expected (De Langhe et al., in press). To our knowledge this is the first report to use proteomics to discover possible genome rearrangements indirectly via the characterization of protein polymorphisms in an allopolyploid crop.

2. Results and discussion

2.1. Characterization of triploid varieties and detection of A and B specific proteins

Previous work on the osmotic stress resistant variety ABB Cachaco (ITC 0643) and the sensitive variety AAA Mbwazirume (ITC 0084) (Carpentier et al., 2007, 2010; Samyn et al., 2007) provided some indications of the existence of inter-variety protein polymorphisms related to their *acuminata* or *balbisiana* origin. In the present study our starting point to detect protein polymorphisms and missing isoforms was to analyze six triploid varieties via classical 2DE: two AAAs (ITC 0081 and ITC 1256), two AABs (ITC 0517 and ITC 0109), one ABB (ITC 0058) and one BBB (ITC 0647). Principal Component Analysis revealed that PC1 could positively be correlated to the presence of the B genome (Fig. 1). Protein spots with a high positive PC1 loading score were abundantly present in the *balbisiana* containing varieties (BBB and ABB) while spots with a high negative PC1 loading score were abundantly present in the *acuminata* containing varieties (AAA). Hierarchical clustering confirmed the PCA results. The first level of clustering separates the BBB and ABB varieties from both AAB and both AAA varieties, the second level separates both AAB varieties from the two AAA varieties and the third level separates both AAA varieties and both AAB varieties demonstrating the proof of principle (Fig. 2).

Download English Version:

https://daneshyari.com/en/article/5165994

Download Persian Version:

https://daneshyari.com/article/5165994

Daneshyari.com