Available online at www.sciencedirect.com
ScienceDirect
Journal of Chromatography A, 1141 (2007) 41–49
JOURNAL OF CHROMATOGRAPHY A
www.elsevier.com/locate/chroma

Fast data preprocessing for chromatographic fingerprints of tomato cell wall polysaccharides using chemometric methods

Bernard Quéméner [a], Dominique Bertrand [b], Isabelle Marty [c], Mathilde Causse [d], Marc Lahaye [a,*]

[a] *INRA, UR Biopolymères, Interactions, Assemblages, rue de la Géraudière, BP 71627, 44316 Nantes Cedex 03, France*
[b] *UMR INRA-ENITIAA, Sensométrie et Chimiométrie, rue de la Géraudière, BP 71627, 44316 Nantes Cedex 03, France*
[c] *INRA, UMR Sécurité et Qualité des Produits d'Origine Végétale, Domaine Saint Paul, Site Agroparc, 84914 Avignon Cedex 9, France*
[d] *INRA, UR Génétique et Amélioration des Fruits et Légumes, Domaine Saint-Maurice, BP 94, 84143 Montfavet Cedex, France*

Received 7 July 2006; received in revised form 21 November 2006; accepted 23 November 2006

Abstract

The variability in the chemistry of cell wall polysaccharides in pericarp tissue of red-ripe tomato fruit (*Solanum lycopersicon* Mill.) was characterized by chemical methods and enzymatic degradations coupled to high performance anion exchange chromatography (HPAEC) and mass spectrometry analysis. Large fruited line, Levovil (LEV) carrying introgressed chromosome fragments from a cherry tomato line Cervil (CER) on chromosomes 4 (LC4), 9 (LC9), or on chromosomes 1, 2, 4 and 9 (LCX) and containing quantitative trait loci (QTLs) for texture traits, was studied. In order to differentiate cell wall polysaccharide modifications in the tomato fruit collection by multivariate analysis, chromatograms were corrected for baseline drift and shift of the component elution time using an approach derived from image analysis and mathematical morphology. The baseline was first corrected by using a "moving window" approach while the peak-matching method developed was based upon location of peaks as local maxima within a window of a definite size. The fast chromatographic data preprocessing proposed was a prerequisite for the different chemometric treatments, such as variance and principal component analysis applied herein to the analysis. Applied to the tomato collection, the combined enzymatic degradations and HPAEC analyses revealed that the firm LCX and CER genotypes showed a higher proportion of glucuronoxylans and pectic arabinan side chains while the mealy LC9 genotype demonstrated the highest content of pectic galactan side chains. QTLs on tomato chromosomes 1, 2, 4 and 9 contain important genes controlling glucuronoxylan and pectic neutral side chains biosynthesis and/or metabolism.

Keywords: Tomato; *Solanum lycopersicon*; Cell wall polysaccharides; QTLs; Texture; HPAEC; Chromatographic data preprocessing; Principal component analysis; Variance analysis

1. Introduction

Fruit texture is an important quality trait determining consumer preferences, post-harvest itineraries up to the retailer shelf and disease resistance [1,2]. Fruit texture results from multiple factors at different scales in which cell walls play essential roles [3]. Biochemical and physico-chemical modifications of cell walls during fruit ripening have been extensively reported [2,4–6] but the intricate and highly regulated mechanisms of cell wall construction [7] and de-assembly [8] in the different cell wall sub-domains (middle lamella, cell wall corners, primary wall) remain only partially known [9,10]. Fleshy fruit cell walls are made of three interconnected networks of cellulose–hemicelluloses, pectins and proteins [11]. Pectins consist of chains of α (1–4)-linked D-galacturonic acid units partially methylesterified and interrupted in "hairy regions" by L-rhamnose carrying arabinan and galactan side-chains [12]. Fruit softening involves a marked reduction in methyl-esterification of pectins, degradation of galactan and arabinan side chains [4,5] and de-assembly of the hemicelluloses–cellulose network [13,14]. Elucidation of these multiple events greatly benefits from the discovery and characterization of texture mutants [15,16] and from the identification of relevant chromosome regions bearing key genes (QTL) related to texture [17,18]. QTLs for physical and sensory attributes of texture have been identified in the progeny of a cherry tomato (*Solanum*

* Corresponding author. Tel.: +33 2 40 67 50 63.
E-mail address: lahaye@nantes.inra.fr (M. Lahaye).

0021-9673/$ – see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.chroma.2006.11.069

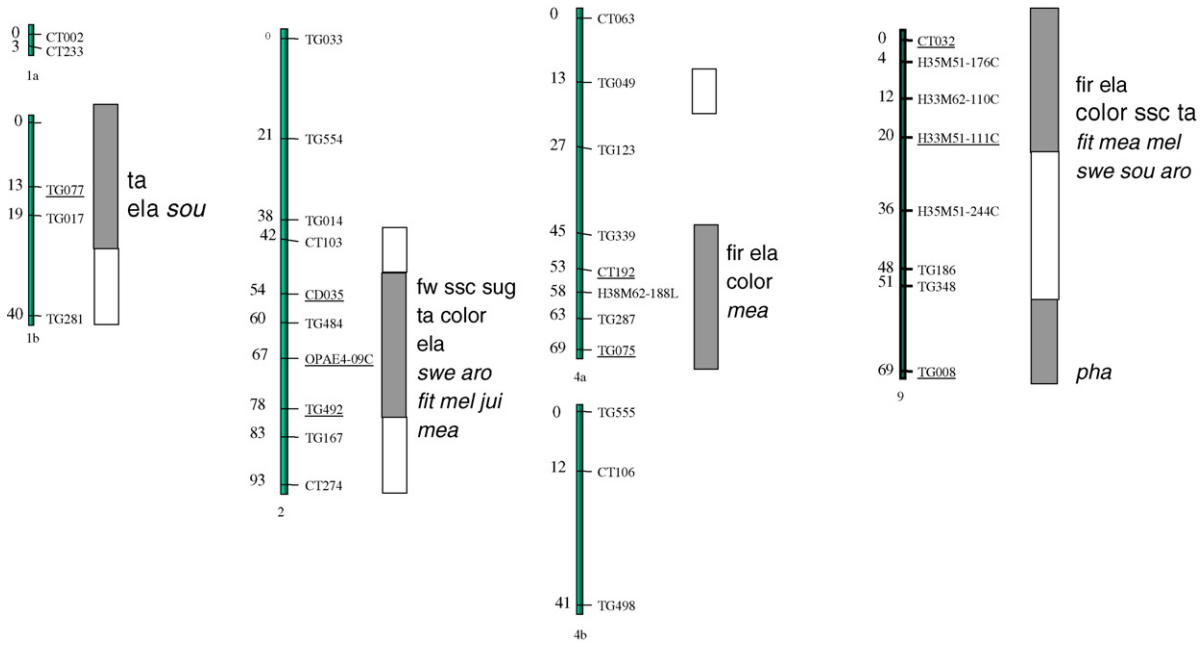

Fig. 1. Genetic map showing the five regions of interest introgressed in LCX and in the QTL–NILs and the molecular markers used to control these regions. Grey boxes correspond to the regions chosen to be introgressed, white boxes indicate the regions which were introgressed with the QTL regions. The traits for which QTLs were detected are indicated on the right of the chromosomes. They corresponded to physical or chemical traits (fw: fruit weight; fir: firmness; ela: elasticity; ta: titratable acidity; ssc: soluble solid content; sug: sugar content; color: fruit color intensity; adapted from ref. [20]) and to sensory traits (swe: sweetness; sou: sourness; aro: overall aroma intensity; fit: firm texture; mel: meltiness; jui: juiciness; mea: mealiness, as described in ref. [19]).

lycopersicon var. cerasiforme (Dun.) Gray cultivar Cervil) and a commercial-type tomato (cultivar Levovil) (Fig. 1; [18–20]). QTLs for firmness were located on chromosomes 2, 5 and 9, the latter explaining 41% of the phenotypic variability. The Cervil alleles on these chromosomes contributed to increase firmness. QTLs for mealiness were located on chromosomes 2–4 and 9 [19]. In order to characterize the impact of these QTLs on cell wall construction, nearly isogenic lines of Levovil tomato carrying Cervil alleles on chromosomes 4 (LC4), 9 (LC9) or on the four chromosomes 1, 2, 4 and 9 (LCX) were grown. We now report on the chemical composition and structure of cell wall polysaccharides of pericarp tissue from the collection of texture-contrasted fruits harvested at the red stage over 2 years. As fine structural analysis of cell wall polysaccharides generally requires enzymatic degradation into oligosaccharides, high performance anion exchange chromatography (HPAEC) known to be a useful technique for oligosaccharide separation [21] was used to obtain informative cell wall polysaccharides enzymatic fingerprints. This analytical approach was recently found useful beside electrophoresis or MALDI–TOF–MS for the analysis of oligosaccharides in the screening for cell wall polysaccharide structural variations in collections of the model plant *Arabidopsis thaliana* [22]. Combined with multivariate analyses of the chromatograms, such enzymatic fingerprints successfully discriminated wheat grain genotypes [23].

HPAEC on pellicular-resin columns is generally carried out in the presence of high-alkaline-pH eluents taking advantage of the electrocatalytic oxydation of sugars at the gold electrode of a pulsed amperometric detector (PAD). In many cases, as far as acidic oligosaccharides or polysaccharides are concerned, sodium acetate (up to 0.5 M) is used in addition. In these conditions, a baseline drift may be observed during the acetate gradient useful for eluting the more strongly retained components. The reasons may be impurities in the sodium acetate used, presence of dissolved oxygen in the eluent and gradual increase in the true electrode area caused by surface reconstruction under the repeated conditions of the oxide on/off cycles in the triply pulsed amperometric waveform [24]. The other drawback of HPAEC is the slow absorption of carbon dioxyde in dilute solution of sodium hydroxyde used through the long runs necessary for eluting acidic oligosaccharides. Carbonate produced acts as a displacing anion causing gradually decreasing retention times. Both of these drawbacks are not critical in usual short analyses, such as for sugar quantification with known standards. But in the case of long chromatographic runs generally as required for a large-scale enzymatic fingerprinting project, the effect of carbonate anions on retention times is still relevant. Malmquist and Danielsson highlighted the sensitivity of principal component analysis (PCA) to all variations in the data set, namely to variations due to the chromatographic process [25]. Bahowick and Synovec, using classic least-squares regression, demonstrated the effect of retention time precision on quantification of liquid chromatographic data [26]. In this context, a pertinent screening methodology based on the multivariate chemometric analysis of the entire chromatographic profile (i.e. the digitized detector signal), without any prior identification nor peak area quantification, necessitates a prerequisite which is the alignment of chromatographic profiles by correcting retention time shifts and baseline drift.

During the past decades, several approaches have been developed in this field [25,27–37]. These approaches may be roughly divided into two groups. In the first one, the alignment is based