



HAL
open science

Comparison of Tree Architecture Using Tree Edit Distances: Application to Two-Year-Old Apple Hybrids

Vincent Segura, Aida Ouangraoua, Pascal Ferraro, Evelyne Costes

► To cite this version:

Vincent Segura, Aida Ouangraoua, Pascal Ferraro, Evelyne Costes. Comparison of Tree Architecture Using Tree Edit Distances: Application to Two-Year-Old Apple Hybrids. *Euphytica*, 2008, 161 (1-2), pp.155–164. 10.1007/s10681-007-9430-6 . hal-00350109

HAL Id: hal-00350109

<https://hal.science/hal-00350109>

Submitted on 30 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Comparison of tree architecture using tree edit distances: application to 2-year-old apple hybrids

Vincent Segura · Aïda Ouangraoua ·
Pascal Ferraro · Evelyne Costes

Received: 2 November 2006 / Accepted: 4 April 2007 / Published online: 1 May 2007
© Springer Science+Business Media B.V. 2007

Abstract In fruit trees, understanding genetic determinisms of architectural traits is considered as a promising manner to control vegetative development and yield regularity. In this context, our study aimed to classify 2-year-old apple hybrids on the basis of their architectural traits. From a fine phenotyping, trees were described as tree graphs, including topological and geometric information. To evaluate the similarity between trees, comparison methods based on edit operations (substitution, insertion and deletion) were carried out. Distance between two tree graphs was computed by minimising the sum of the costs of the edit operations applied to transform one tree into another. Two algorithms for the comparison of unordered and partially ordered tree graphs were applied to a sub-sample of the population, taking into account several geometric attributes. For each comparison, a dissimilarity matrix was computed, and subsequently trees were clustered. A local interpretation of the matched entities was proposed through schematic representations of the trees, and similarities between trees were analysed within and between

clusters. The tree graphs, both unordered or partially ordered and whether the attributes were considered or not, were grouped, by clustering, according to the number of entities per tree. The robustness of the unordered comparison was demonstrated by its application to the whole population, since it provided results similar to those obtained on the sub-sample. Further developments towards a higher relative weight of geometric versus topological information are discussed in the perspective to define an architectural ideotype in apple.

Keywords Branching · Clustering · Geometry · *Malus x domestica* · Topology · Tree graph

Introduction

The domesticated apple tree (*Malus x domestica* Borkh.) is one of the most important temperate fruit trees. Thus, breeding programmes are carried out in this species to select productive cultivars with high fruit quality and resistance to pests and diseases (Lespinasse 1992; Laurens 1999). However, the consideration of tree architecture and shoot morphology constitutes also a way to select trees adapted to innovative orchard management, by improving the control of vegetative development and yield regularity (Laurens et al. 2000).

A large architectural variability exists between apple cultivars, which have been classified into four

V. Segura (✉) · E. Costes
INRA, UMR DAP, Equipe Architecture et
Fonctionnement des Espèces Fruitières, 2 place Viala,
34060 Montpellier Cedex, France
e-mail: segura@supagro.inra.fr

A. Ouangraoua · P. Ferraro
UMR LaBRI, Université de Bordeaux 1, 351 Cours de la
Libération, 33405 Talence Cedex, France

architectural types by qualitative criteria such as growth habit, distribution of branches and fruiting position (Lespinasse 1977). Over the last 10 years, more detailed architectural studies have been performed, especially through the distinction between topology (i.e. the relative positions of entities within the tree) and geometry (i.e. the spatial position and form of entities) (Godin et al. 1999). Tree habit variability has been quantified in apple tree through the investigation of topological organisation, in particular, by the variability of branching patterns (Lauri et al. 1995; Costes and Guédon 2002; Renton et al. 2006). Tree habit is also a result of the branch form. A bio-mechanical model of apricot branches showed that the main factors involved in the final branch form were its initial geometry (particularly, slenderness and inclination) and the load distribution along the branch (Alméras et al. 2004). Architectural traits are also expected to be related to yield regularity: fructification regularity has been shown to be related to spur extinction (i.e. the death of short shoots) in apple cultivars (Lauri et al. 1995). Both geometric and topological traits are thus putative targets for apple breeding programmes. However, even if architectural variability has been shown and quantified in apple species, the genetic determinisms of architectural traits and their recombination within genotypes remain poorly understood.

In this context, a quantitative genetics approach has been initiated at INRA Montpellier. From a fine phenotyping at a juvenile stage, we have already shown that many geometric and topological traits displayed high heritability and strong genetic control (Segura et al. 2006, 2007). Moreover, to investigate how the architectural traits segregate and recombine within progenies and to contribute to the definition of architectural ideotypes (Lespinasse 1992; Dickman et al. 1994), several apple classifications have already been proposed. De Wit et al. (2004) proposed a classification of 2-year-old apple hybrids on the basis of the hierarchical organisation of the trees, as previously introduced by Edelin (1991). This classification displayed a continuum from trees with a few short branches to trees with many long branches. We have also proposed a classification of 1-year-old apple hybrids on the basis of heritable and uncorrelated traits (Segura et al. 2006). However, in these

studies, the topological organisation of the trees was poorly exploited. A first step in this direction was reached by the comparison of branching sequences along axes of diverse apple cultivars (Costes and Guédon 2002; Guédon et al. 2003). However, that methodology is only applicable to axes and the whole tree topology was not accounted for. To extend this comparison method to entire branching systems, Ferraro and Godin (2000) have adapted a comparison algorithm, initially developed by Zhang (1996). The present work aimed at investigating this comparison technique on a 2-year-old apple F1 progeny. Several comparison techniques were applied to apple hybrids, formalised as tree graphs, taking or not into account topologic and geometrical information. From the resulting dissimilarity matrices, clustering techniques were carried out to highlight groups of hybrids. A local interpretation of the comparisons was proposed through schematic representations of the trees, and similarities between trees were analysed within and between clusters. Clustering variability was also studied with respect to the comparison method applied and the attributes accounted. Prospects given by this approach in a breeding context are discussed.

Materials and methods

Plant material

The progeny under study was derived from the ‘Starkrimson’ × ‘Granny Smith’ cross. Parents were chosen for their contrasted architecture. According to Lespinasse (1992), the Starkrimson maternal parent displayed an erected growth habit, a fructification located on short shoots, and a tendency to irregular bearing. The ‘Granny Smith’ pollen parent is characterised by a weeping habit, a terminal fructification on long shoots and brindles and a tendency to yield regularity. The studied population includes 50 genotypes randomly sampled from the ‘Starkrimson’ × ‘Granny Smith’ progeny initially composed of 125 genotypes. The 50 genotypes were replicated three times by grafting onto ‘Pajam 1’ rootstock. These 150 trees were implanted in March, 2003 at the Melgueil INRA experimental station (Segura et al. 2006).

Phenotyping

About 148 trees were observed in 2005, after the second year of growth. At that time, they consisted of a trunk, sylleptic axillary shoots (i.e. shoots that have grown the same year as their parent shoot), and proleptic axillary shoots (i.e. shoots that have grown 1 year after their parent shoot) (Fig. 1). Three types of axillary shoots were distinguished by their length: long shoots (longer than 20 cm), brindles (between 5 and 20 cm) and spurs, or short shoots (shorter than 5 cm). The topology was recorded using the representation of multiscale tree graphs (Godin and Caraglio 1998) and the coding method defined by Godin et al. (1997). Three scales were defined within the trees: the whole tree (scale 1), the axes (scale 2) and the internodes (scale 3). Tree topology was recorded until scale 2, i.e. all the axes within the tree were noted. Each axis was labelled, according to its length, as ‘L’ for long shoots, ‘B’ for brindles and ‘S’ for spurs. To account for more precise branching information, topology was acquired at scale 3 for the trunk, four long sylleptic axillary shoots (LSAS) and four long proleptic axillary shoots (LPAS), i.e. the

rank of the shoots they born was also integrated. In addition, geometric features were recorded on these sampled axes. The length, bottom and top diameters were measured on the trunk, LSAS and LPAS. The bottom and top angles, cord and projection on the ground were measured on LSAS and LPAS. Measured variables were used to calculate others to provide descriptors as close as possible to biological processes. For example, the conicity was calculated as: $Coni = (D_{bot} - D_{top})/Le$, where D_{bot} and D_{top} are, respectively, the bottom and top diameters, and Le the length of an axis. The orientation was calculated as: $Orient = \alpha (1 - C/P)$, where $\alpha = -1$ when the branch is weeping, $\alpha = 1$ when it is erect, C is the cord and P the projection of the branch.

Tree graph representation of plant architecture

The phenotyping method used to record a plant architecture allowed us to formalise plants as tree graphs (Godin and Caraglio 1998). A plant is viewed as an assembly of adjacent botanical components, such as internodes, annual growth units or axes. Such a structure can be formally described by defining a set

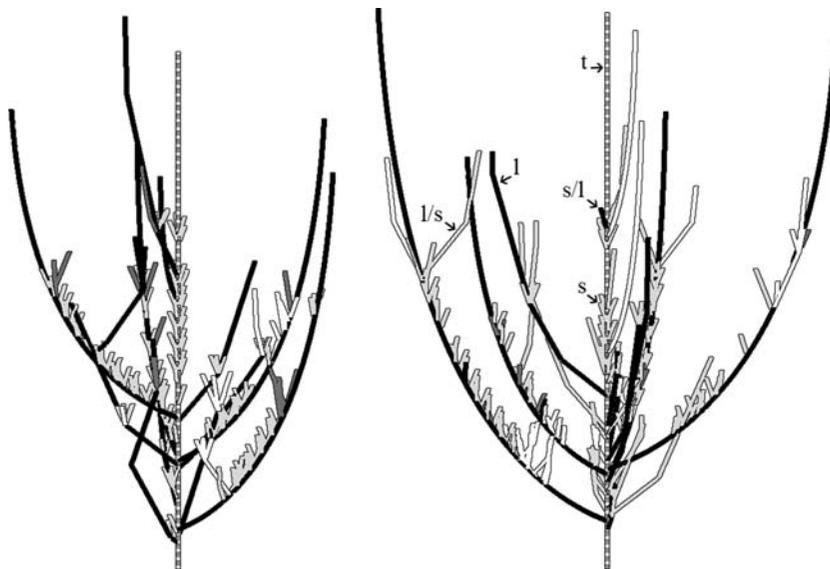


Fig. 1 A schematic representation of an unordered comparison of a pair of trees. Axes in the first tree were coloured according to their length class: the trunk was striped in *grey* and *white*, the long shoots were coloured in *black*, the brindles in *dark grey* and the spurs in *light grey*. Axes in the second tree were coloured according to the length class of the axes they matched in the first tree. Unmatched axes, resulting from insertion and

deletion, were coloured in *white*. Matching between axes belonging to a same class length are labelled with the letters *t*, *l* and *s*, for trunks, long shoots and spurs matched together. Matching between axes from different class length are labelled out by *l/s* and *s/l* for long shoots matched with spurs and spurs with long shoots

of vertices, V , that represent the plant components, and a list, E , of vertex pairs that describes the adjacency of these components. We assume that each component, x , is physically attached to the plant body by at most one parent component, called parent of x . For any pair of vertices (x, y) in E , x and y are a parent–child pair. The resulting topological structure is a rooted tree graph, $T = (V, E)$. In such a graph, every vertex except the root, r , has exactly one parent vertex. Three kinds of tree graphs can be distinguished. A tree on which an order is imposed, by assigning an ordering to the children of each vertex, is an ordered tree. Ouangraoua et al. (submitted) recently introduced the notion of a partially ordered tree. A partially ordered tree is one where some children are ordered on their parent vertex, while others are not. An unordered tree is a tree on which there is no ordering imposed.

In our data set, since topological notations at scale 3 were only recorded for the trunk, four LSAS and four LPAS, the tree graphs were partially ordered. The sampled axes were ordered, the rest of the tree was unordered. Here, the ordering is the rank of the axes born by the sampled axes. Tree graphs were also considered as unordered tree graphs. In that case, only the presence and the number of shoots were considered. Comparisons of the tree graphs were performed at the scale of axes (scale 2) using two comparison algorithms, depending on whether the graphs were unordered or partially ordered.

Edit distance between tree graphs

Zhang (1996) and Ouangraoua et al. (submitted) have proposed two algorithms for computing an edit distance between unordered tree graphs and between partially ordered tree graphs, respectively. The edit distance between two trees, T_1 and T_2 , is the minimum cost of the elementary operations needed to transform T_1 into T_2 . Three edit operations are considered: substituting one vertex into another, deleting (i.e. removing a vertex v from T_1 , and making its children become the children of its father) or inserting one vertex (i.e. the symmetric operation on T_2). To transform one tree graph into the other, all the vertices of T_1 and T_2 must be affected by at most one edit operation. A cost function, called the local distance, is defined for each edit operation, s . The local distance maps a non-negative real number $\gamma(s)$

to s . $\gamma(s) = d_{\text{sub}}(v_1, v_2)$ if s substitutes the vertex v_1 into the vertex v_2 , $\gamma(s) = d_{\text{del}}(v_1) = d(v_1, \lambda)$ if s deletes the vertex v_1 , and $\gamma(s) = d_{\text{ins}}(v_2) = d(\lambda, v_2)$ if s inserts the vertex v_2 . The cost of edit operations depends on the use of axis attributes in the comparison, such as the length class or geometric measures. When no attributes are included in the comparison, the substitution is called a match, its cost is zero and the insertion–deletion cost is one. When attributes are included, the substitution cost remains zero if the substituted entities have the same attributes. In this case the substitution is a match. Otherwise, the cost is positive and its value can be tuned by a user according to his biological objectives.

The dissimilarity measure, $D(T_1, T_2)$, from a tree graph T_1 to a tree graph T_2 , is the minimum cost of transforming T_1 into T_2 : $D(T_1, T_2) = \min\{\sum_{s \in S} \gamma(s), \text{ such that } S \text{ is a sequence of edit operations that takes } T_1 \text{ to } T_2\}$. The notion of mapping is used to characterise a sequence of edit operations that transforms T_1 into T_2 , ignoring the order in which they are applied. A mapping M , is a set of ordered pairs of vertices (v_1, v_2) from $T_1 \times T_2$, where v_1 and v_2 are substituted during the sequence of edit operations. Zhang (1996) and Ouangraoua et al. (submitted) have shown that a mapping must respect two constraints to be equivalent to a sequence of edit operations: a one-to-one constraint, i.e. one axis in the first tree is substituted by at most one axis in the second, and a structure preservation constraint, i.e. two axes in a branching system are not substituted by two axes in two different branching systems. These two constraints are common to both comparison methods. In the case of partially ordered tree graphs comparison, there is an additional constraint, the preservation of the partial order on the axes, i.e. accounting for the rank of the lateral shoots in the tree comparison. Zhang (1996) and Ouangraoua et al. (submitted) proposed two polynomial algorithms for the unordered and the partially ordered cases, respectively, to compute the optimal mapping and measure the edit distance between two trees.

The mapping corresponding to the optimal sequence of edit operations between the plant entities was visualised using a three-dimensional (3D) plant reconstruction. Colouring tools used for the 3D representation provided feed back on the detailed matching between elementary tree entities, and allowed an analysis of the comparisons between tree

pairs. Distance matrices were also computed from the comparisons. To highlight groups of trees, clustering was applied to distance matrices using the Ward's method (Ward 1963).

Comparisons and 3D visualisations were performed using, respectively, the TreeMatching function and the GEOM module of AMAPmod software (Godin and Guédon 2003), and clustering was done using proc Cluster and proc Tree in the SAS software (SAS Institute Inc 2000).

To investigate and compare the two comparison methods on the basis of a comprehensive interpretation of matched entities between trees, they were first applied to a sub-sample of hybrids. This sub-sample was selected from a partition of the whole population obtained from the following heritable and uncorrelated quantitative traits: number of axes per tree, mean diameter of the second annual growth of the trunk, trunk mean internode length, LPAS conicity and LSAS orientation (Segura et al. 2006). On the basis of these traits, hybrids were clustered into six groups and the three trees at the centre of each group were selected to constitute a sub-sample of 18 trees. First, comparisons of unordered tree graphs were performed on this sub-sample, both including and not including the entity attributes such as their length class or geometric measures (geometric traits from which the sub-sample was selected). Second, comparisons of unordered tree graphs, without attributes, were applied to the whole progeny. Third, partially ordered comparisons were applied to the 18 tree sub-sample.

Results

Unordered comparison

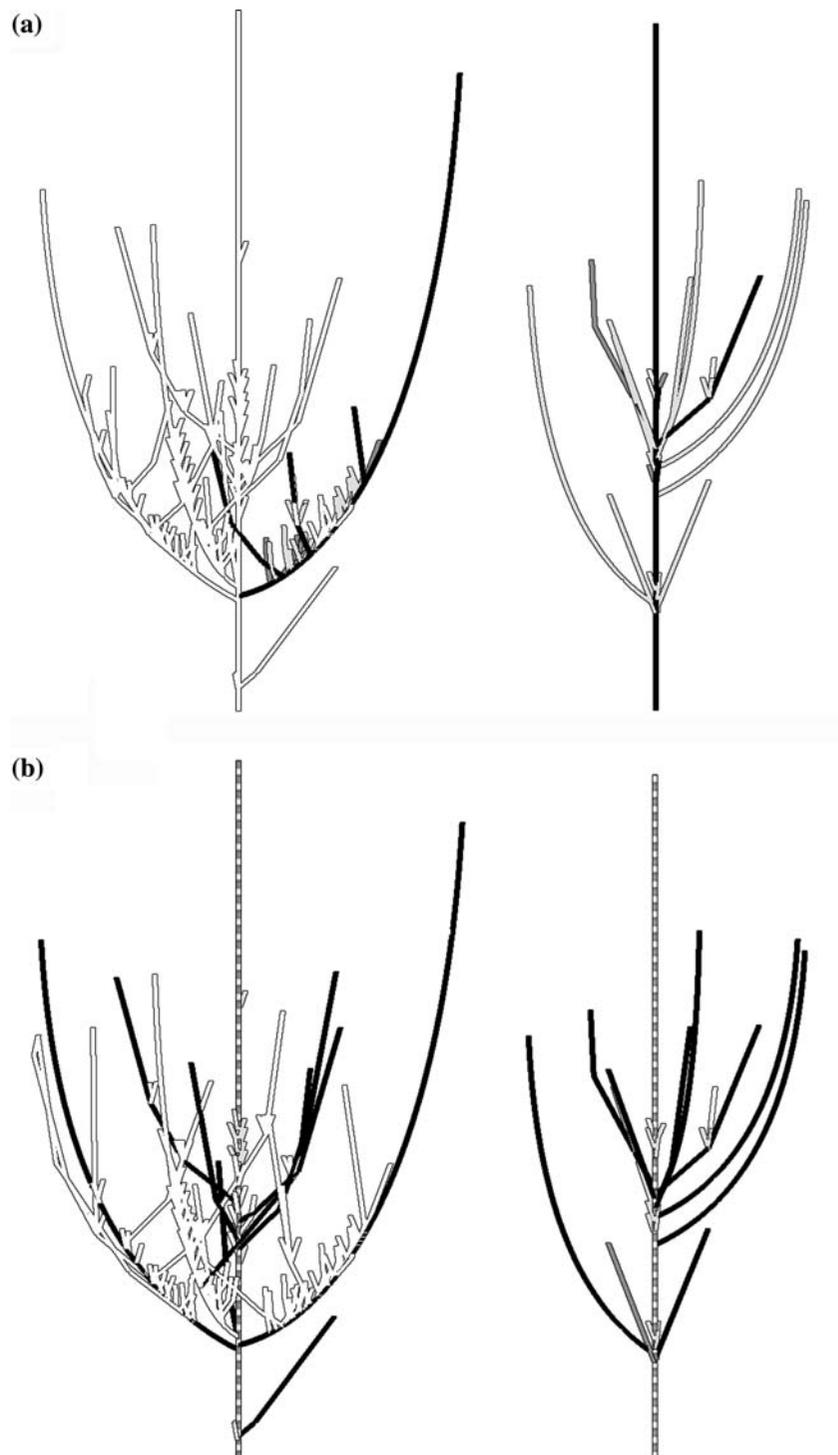
Unordered comparisons have first been analysed through 3D visualisations (Fig. 1). For visualisation purposes, the unordered graphs were schematised as partially ordered graphs. Similarly, axes in the first tree were coloured according to their length class even though comparison did not account for this information: the trunk was striped in grey and white, long shoots were coloured in black, brindles in dark grey and spurs in light grey. Axes in the second tree were coloured according to the length class of the axes in the first tree they substituted or matched.

Unmatched axes, resulting from insertion and deletion, were coloured in white. This representation highlighted that in many cases, axes belonging to the same length class were matched. In Fig. 1, on the second tree, some axes matched to axes on the first tree are labelled. The matching trunk is labelled with 't', the long shoot with 'l' and the spur with 's'. Long shoot matched to a spur is labelled with 'l/s' and spur matched to a long shoot with 's/l'. In the comparison shown in Fig. 2a, a trunk was matched with a LSAS. A match of a trunk to a LSAS was not satisfactory and led us to take into account the length class as axis attribute in further comparisons. In the last example, shown in Fig. 2b, accounting for the length class resulted in trunks that were properly matched. This correction increased the distance between the trees, from 0.69 to 2. Thus, accounting for entity length class attributes modified the mapping during the comparisons between unordered tree pairs.

Unordered comparisons were also analysed using Ward's clustering method on the dissimilarity matrix. The cluster dendrogram, performed on the comparison of the 18 trees forming the sub-sample, displayed three clusters according to the decrease in semi-partial R^2 (Fig. 3). Further analysis of the group composition showed that differences between groups were related to the total number of axillary shoots per tree (Table 1). Group 1 was composed of trees with an intermediate number of shoots (on average 104.0 shoots), group 2 was composed of trees with many shoots (on average 216.6 shoots) and group 3 was composed of sparsely branched trees (on average 32.0 shoots). Moreover, the clustering remained almost the same when class length or geometric attributes were included in the unordered comparison (data not shown). In fact the number of insertions and deletions remained similar with or without considering attributes. This shows that the number of entities at the scale of comparison is preponderant in the edit distance.

An unordered comparison was also applied to the whole population. Again three groups were displayed according to the decrease in semi-partial R^2 (data not shown). As well as for the sub-sample comparison, differences between the three groups were related to the branching amount, and the average number of axes per tree in each group was very similar to that observed in the sub-sample comparison (Table 1). Moreover, 17 of the 18 trees were in the same groups

Fig. 2 Schematic representation of an unordered comparison of a pair of trees, **(a)** without axes attributes, and **(b)** with length class accounted as attributes in the comparison. For axis colour explanations, see Fig. 1



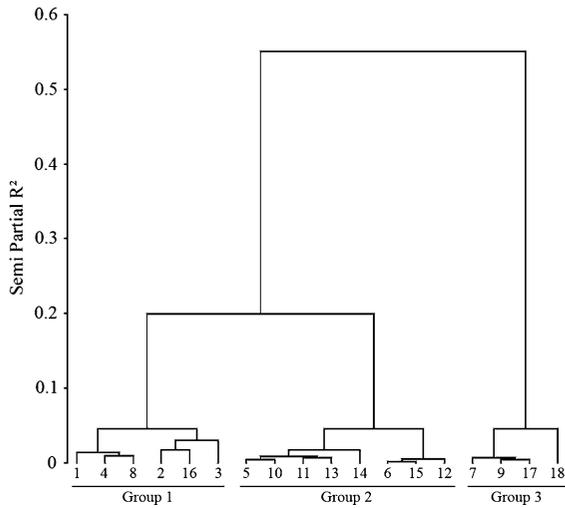


Fig. 3 Cluster dendrogram from the distance matrix resulting from the unordered comparison of a sub-sample of 18 apple hybrids. Dissimilarity, indicated by the vertical axis, corresponds to the between cluster semi-partial R^2 . Hybrid numbers are on the leaves of dendrogram. Groups resulting from the selected separating level and used in the analysis are indicated as Groups 1–3

Table 1 Average number of axes per tree for the groups 1–3 of the clusterings performed from the unordered comparison, either on a sub-sample of 18 trees or on the whole apple progeny

	Group 1	Group 2	Group 3
Sub-sample	104.0 (39.6)	216.6 (36.4)	32.0 (13.2)
Whole progeny	94.7 (29.0)	224.0 (78.7)	37.8 (16.4)

Standard deviations are indicated into brackets

as those obtained from the previous comparisons of the sub-sample.

Partially ordered comparison

Partially ordered comparisons of the 18 tree sub-sample were also performed. Analysis of schematised tree pairs was not suitable for comparing this method with the unordered comparison because axes could not be coloured according to their order as they were coloured according to their class length. In contrast, both methods were compared through the analysis of clustering variability. The clustering was slightly different than that of the unordered tree graphs (Fig. 4). Because of the decrease in semi-partial R^2 , two separation levels were considered. In the first

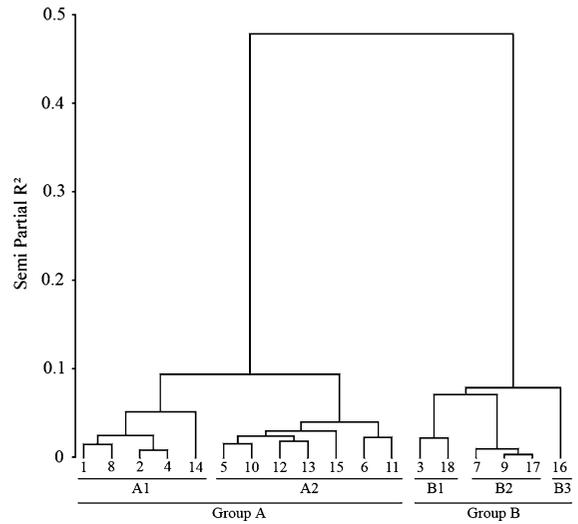


Fig. 4 Cluster dendrogram from the distance matrix resulting from the partially ordered comparison of a sub-sample of 18 apple hybrids. Dissimilarity, indicated by the vertical axis, corresponds to the between cluster semi-partial R^2 . Hybrid numbers are on the leaves of dendrogram. Groups and sub-groups resulting from the selected separating level and used in the analysis are indicated as groups A (A1 and A2) and B (B1–B3)

level, there were two groups, A and B. As for the unordered comparison, these two groups differed in the number of shoots per tree. Trees belonging to group A had an average of 185.3 shoots, and trees in group B had an average of 43.7 shoots. In the second level, five groups were distinguished. Group A was divided into two groups, A1 and A2, that were discriminated by the number of shoots per tree. They are similar to the groups of intermediate (group 1) and highly branched trees (group 2) of the unordered comparisons, respectively. Group B, composed of sparsely branched trees, was divided into three groups. Differences between the groups were not only the number of shoots per tree, but also the axis positions along the trunk (Table 2). Differences in the number of shoots isolated the group B3, which was composed of only one tree, and discriminated the groups B1 and B2. But trees of B1 and B2 also differed in the rank of the last axillary shoot of the trunk and in the length of the branching zone of the trunk. For the groups of intermediate (A1) and highly branched trees (A2), the influence of axes position was detected deeper in the clustering (data not shown). Thus, partially ordered comparisons allowed

Table 2 Average values of traits characterising the position of lateral axes along the trunk for the groups B1–B3 performed from the partially ordered comparison on a sub-sample of 18 apple hybrids

Traits	Group B1	Group B2	Group B3
Number of axillary shoots	55.5 (7.8)	24.7 (4.7)	71.0
First axillary shoot rank	11 (4.2)	15 (8.7)	11
Last axillary shoot rank	115 (2.8)	71 (1.0)	71
Branching zone length	95 (1.4)	57 (8.7)	61

Standard deviations are indicated into brackets and were not calculated for the group B3 because it was composed of only one tree

us to cluster hybrids firstly by the number of shoots per tree and secondly by the axes positions. This second criterion had a visible weight particularly when the compared trees had a similar number of entities and when the number of entities per tree was small.

Discussion

This study aimed at investigating the architectural variability of apple hybrids by comparison methods. The phenotyping method used allowed us to describe apple trees as tree graphs, that were used to perform comparisons of trees. This approach is a step forwards in the classification of branching structures in apple over previous studies, which classified apple hybrids on the basis of quantitative traits (De Wit et al. 2004; Segura et al. 2006). The approach proposed in the present paper accounts for the entire tree and thus extends more local structural comparisons, such as those provided by branching sequences alignment (Guédon et al. 2003).

The unordered comparison was first carried out on a sub-sample and groups obtained from clustering were related to the number of entities at the comparison scale, as already shown by Ferraro and Godin (2000). In that study, the authors recommended that comparison should only be made to plants with similar numbers of entities. However, in the present work, trees varied widely in the number of shoots and thus the three groups of hybrids were clearly separated by the number of entities per tree. Moreover, this clear separation in three groups, in relation with the number of shoots, was confirmed in the comparison of the whole population, underlining the robustness of the unordered comparison method. This application to a large number of hybrids is a step

in the use of comparison methods in a genetic context, i.e. where studied populations are usually large. In this perspective, it would be of interest to link phenotypic and genotypic variability. This could be performed on the whole progeny, which contained 125 genotypes in total, and for which both phenotypic data and genetic map are available. Moreover, QTLs were mapped for branching traits, such as the total number of sylleptic axillary shoots developed during the first year of growth (Segura et al. 2007). A possible extension of the present study could consist in comparing and clustering the whole progeny and linking these results with QTL analysis. In particular, the combination of alleles from QTLs detected for sylleptic branching could be screened between the clusters.

To further develop our analysis, another comparison method of partially ordered tree graphs was carried out (A. Ouangraoua et al., submitted data). The use of this technique was motivated by the possibility of accounting for axes position on their parent axes, especially on the trunk. There is much interest in the hierarchical organisation of tree, as relating it to its habit (De Wit et al. 2004). Weeping apple trees are usually characterised by an acrotonic organisation, long shoots are near the top of the trunk, and erected trees are often characterised by a basitonic organisation, long shoots are near the bottom of the trunk (Lespinasse 1992). Additionally, partially ordered comparisons provide a new method that is well adapted to data sets where the trees are partially described with a sampling procedure. This is often the case in genetic studies where a phenotyping strategy is used to screen a large number of individuals. A large number of genotypes is required for QTL mapping (Beavis 1998), and replicates of genotypes are also suitable for broad sense heritability estimates (Gallais 1989). Partially ordered

comparisons were carried out on the sub-sample and these were compared to the unordered comparisons of the sub-sample. The clustering resulted again from the number of entities at the scale of comparison, but slight modifications were also observed. These modifications were related first to the number of clusters and second to the variables that discriminate the clusters. Indeed, differences between clusters were not only the number of shoots per tree but also the axes position along the trunk, especially for sparsely branched trees. Even though the relative weight of axes positions was low in the clustering, such results are promising with regard to our initial objective to integrate the information of the axes position in the tree comparisons. In the next future, the partially ordered comparison should be tested on the whole population to test its robustness.

Additionally, in the objective to contribute to the definition of architectural ideotypes in apple tree, entity attributes were accounted in the tree comparisons. The concept of ideotype was first proposed by Donald (1968) and further developed in trees by Dickman et al. (1994). According to this author, to establish a tree ideotype one needs to specify the priority traits that may be associated in a given tree according to the initial objectives. Considering architectural traits in apple species, objectives are typically related to yield regularity, and control of the tree growth and habit for orchard management (Lauri and Costes 2004). For reaching this last objective, topological and geometric traits are putative targets, which may be combined in selected genotypes. As seen before, tree habit results from the hierarchical organisation of branching along the trunk, and from the flexion of branches (Alméras et al. 2004). Different attributes were thus included in tree comparisons to explore a new possibility for analysing combinations of suitable traits in hybrids. The attributes considered here were (1) axes length class, assumed to be involved in the hierarchical organisation of the tree; (2) local geometric traits characterising tree growth (internode length and trunk diameter); and (3) geometric traits characterising initial branch form and position (conicity and orientation). However, when these attributes were considered in the unordered comparison, even if an effect was detected through the local analyses, no difference was detected in the final clustering. This mainly resulted from the preponderant weight of insertion

and deletion operations in the distance computation. In the next future it would be thus of interest to balance the relative weight of topology and attributes. One approach that could be used to address this issue may consist in comparisons without account for the extremities of tree graphs. Such technique, called extremity space-free, has already been applied to branching sequences and successfully decreased the weight of the sequence length in the comparison (Guédon et al. 2003). Using this new method would allow comparisons of sub-trees with a comparable number of entities as recommended by Ferraro and Godin (2000). This new method has from now been implemented in unordered tree graph comparison and is going to be tested on our data as a continuation of the present work.

Acknowledgements This research was partly funded by a Ph.D. grant from ‘ACI Arborescences’ project, allocated to A. Ouangraoua, and a Ph.D. grant from INRA Genetic and Breeding Department and Languedoc-Roussillon Region, allocated to V. Segura. We are grateful to all the members of the ‘ACI Arborescences’ project for their fruitful discussions. We acknowledge G. Garcia and S. Feral for their contribution to field measurements and technical assistance in the orchard. We also acknowledge C. Smith for improving the English.

References

- Alméras T, Costes E, Salles JC (2004) Identification of bio-mechanical factors involved in stem shape variability between apricot tree varieties. *Ann Bot* 93:455–468
- Beavis WD (1998) QTL analyses: power, precision, and accuracy. In: Paterson AH (ed) *Molecular dissection of complex traits*. CRC, New York
- Costes E, Guédon Y (2002) Modelling branching patterns in 1-year-old trunks of six apple cultivars. *Ann Bot* 89:513–524
- De Wit I, Cook NC, Keulemans J (2004) Characterization of tree architecture in two-year-old apple seedling populations of different progenies with a common columnar gene parent. *Acta Hort* 663:363–368
- Dickman DI, Gold MA, Flore JA (1994) The ideotype concept and the genetic improvement of tree crops. *Plant Breed Rev* 12:163–193
- Donald CM (1968) The breeding of crop ideotypes. *Euphytica* 17:385–403
- Edelin C (1991) Nouvelles données sur l’architecture des arbres sympodiaux: le concept de plan d’organisation. In: Edelin C (ed) *L’Arbre, Biologie et développement*. Naturalia Monspeliensia, Montpellier
- Ferraro P, Godin C (2000) A distance measure between plant architectures. *Ann For Sci* 57:445–461
- Gallais A (1989) *Théorie de la sélection en amélioration des plantes*. Masson, Paris

- Godin C, Caraglio Y (1998) A multiscale model of plant topological structures. *J Theor Biol* 191:1–46
- Godin C, Costes E, Sinoquet H (1999) A method for describing plant architecture which integrates topology and geometry. *Ann Bot* 84:343–357
- Godin C, Guédon Y (2003) AMAPmod: exploring and modeling plant architecture. Version 2.1.16. CIRAD/INRA UMR Modélisation des plantes, Montpellier
- Godin C, Guédon Y, Costes E et al (1997) Measuring and analyzing plants with the AMAPmod software. In: Michalewicz M (ed) *Advances in computational life science*. CSIRO, Australia
- Guédon Y, Heuret P, Costes E (2003) Comparison methods for branching and axillary flowering sequences. *J Theor Biol* 225:301–325
- Laurens F (1999) Review of the current apple breeding programmes in the world: objectives for scion cultivar improvement. *Acta Hort* 484:163–170
- Laurens F, Audergon J, Claverie J et al (2000) Integration of architectural types in French programmes of ligneous fruit species genetic improvement. *Fruits* 55:141–152
- Lauri PE, Costes E (2004) Progress in whole tree architectural studies for apple cultivar characterization at INRA, France—contribution to the ideotype approach. *Acta Hort* 663:357–362
- Lauri PE, Terouanne E, Lespinasse JM et al (1995) Genotypic differences in the axillary bud growth and fruiting pattern of apple fruiting branches over several years—an approach to regulation of fruit bearing. *Sci Hort* 64:265–281
- Lespinasse JM (1977) La conduite du pommier: types de fructification, incidence sur la conduite de l'arbre. INV-UFLEC, Paris
- Lespinasse Y (1992) Le pommier. In: Gallais A, Bannerot H (eds) *Amélioration des espèces végétales cultivées—objectifs et critères de sélection*. INRA Editions, Paris
- Renton M, Guédon Y, Godin C et al (2006) Similarities and gradients in growth unit branching patterns during ontogeny in 'Fuji' apple trees: a stochastic approach. *J Exp Bot* 57:3131–3143
- SAS Institute Inc (2000) SAS user's guide: statistics. SAS Institute Inc., Cary, NC, USA
- Segura V, Cilas C, Laurens F et al (2006) Phenotyping progenies for complex architectural traits: a strategy for 1-year-old apple trees (*Malus x domestica* Borkh.). *Tree Genet Genome* 2:140–151
- Segura V, Denancé C, Durel CE et al (2007) Wide range QTL analysis for complex architectural traits in a 1-year-old apple progeny. *Genome* 50:159–171
- Ward JH (1963) Hierarchical grouping to optimize an objective function. *J Am Stat Assoc* 58:236–244
- Zhang K (1996) A constrained edit distance between unordered labeled trees. *Algorithmica* 15:205–222