Root System Markup Language: toward a unified root architecture description language

Guillaume Lobet, Michael P. Pound, Julien Diener, Christophe Pradal, Xavier Draye, Christophe Godin, Mathieu Javaux, Daniel Leitner, Félicien Meunier, Philippe Nacry, et al.

▶ **To cite this version:**

Guillaume Lobet, Michael P. Pound, Julien Diener, Christophe Pradal, Xavier Draye, et al.. Root System Markup Language: toward a unified root architecture description language. Plant Physiology, 2015, 167 (3), pp.617-627. 10.1104/pp.114.253625 . hal-01113767

HAL Id: hal-01113767
https://hal.science/hal-01113767

Submitted on 6 Feb 2015

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.

Plant Physiology Preview. Published on January 22, 2015, as DOI:10.1104/pp.114.253625

Running head: Unified root system architecture language

Xavier Draye
Croix du Sud, 2-L7.05.11
1348 Louvain-la-Neuve
Belgium

Tel : +32 10 47 20 92
e-mail : xavier.draye@uclouvain.be

Journal research area:
Breakthrough technology

Copyright 2015 by the American Society of Plant Biologists

Root System Markup Language: toward an unified root architecture description language

Guillaume Lobet[1]*, Michael P Pound[2]*, Julien Diener[3]*, Christophe Pradal[3,4]*, Xavier Draye[5†], Christophe Godin[3], Mathieu Javaux[5,9], Daniel Leitner[6], Félicien Meunier[5], Philippe Nacry[7], Tony P Pridmore[8], Andrea Schnepf[9].

[1.] PhytoSYSTEMS, Université de Liège, Boulevard du Rectorat, 27, 4000 Liège, Belgium
[2.] Centre for Plant Integrative Biology, School of Biosciences, University of Nottingham, Sutton Bonington, LE12 5RD, UK.
[3.] Virtual Plants, INRIA, CIRAD, INRA, 34095 Montpellier France
[4.] Institut de Biologie Computationnelle, F-34095 Montpellier, France
[5.] Earth and Life Institute, Université catholique de Louvain, Croix du Sud 2 – L7.05.11, 1348, Louvain-la-Neuve, Belgium
[6.] Computational Science Center, University of Vienna, Oskar Morgenstern-Platz 1, 1090 Vienna, Austria
[7.] Biochemistry and Plant Molecular Physiology, UMR 5004 CNRS/INRA/SupAgro-M/UM2. Institut de Biologie Intégrative des Plantes. Place Viala, 34060 Montpellier Cedex 1, France.
[8.] School of Computer Science, University of Nottingham, Jubilee Campus, Wollaton Rd, Nottingham, NG8 1BB, UK
[9.] Institut für Bio- und Geowissenschaften : Agrosphäre (IBG-3), Forschungszentrum Jülich GmbH, D-52425 Jülich, Germany

* These authors contributed equally to this work
† Corresponding author

Note: The first four authors contributed equally to this work. Others were ordered alphabetically.

One-Sentence summary:
Introducing a unified format for root system architecture description

Footnotes :

This work was supported by the Belgian Science Policy Interuniversity Attraction Poles Program (grant P7/29 to G.L., X.D), by the Fonds National de la Recherche Scientifique (G.L. and F.M.), by the French National Research Agency (HydroRoot project, grant ANR-11-BSV6-0018 to C.P., C.G.), by EU FP7-KBBE-2011-5 project (EURoot: Enhancing resource Uptake from roots under Stress in Cereal Crops to P.N., T.P.P., M.P.P and X.D.), by the Agropolis Foundation Montpellier France (Rhizopolis grant to P.N., C.P., J.D. and C.G.), by the German research association (DFG, Transregio Collaborative Research Center 32, Patterns in Soil–Vegetation–Atmosphere Systems: Monitoring, Modelling and Data Assimilation to A.S.) and by the Austrian Academy of Sciences (APART-fellowship of the at the Computational Science Center at the University of Vienna to D.L.).

Corresponding author :

xavier.draye@uclouvain.be

Abstract

The number of image analysis tools supporting the extraction of architectural features of root systems has increased over the last years. These tools offer a handy set of complementary facilities, yet it is widely accepted that none of these software tool is able to extract in an efficient way growing array of static and dynamic features for different types of images and species. .

We describe the Root System Markup Language (RSML) that has been designed to overcome two major challenges: (i) to enable portability of root architecture data between different software tools in an easy and interoperable manner allowing seamless collaborative work, and (ii) to provide a standard format upon which to base central repositories which will soon arise following the expanding worldwide root phenotyping effort.

RSML follows the XML standard to store 2D or 3D image metadata, plant and root properties and geometries, continuous functions along individual root paths and a suite of annotations at the image, plant or root scales, at one or several time points. Plant ontologies are used to describe botanical entities that are relevant at the scale of root system architecture. An xml-schema describes the features and constraints of RSML and open-source packages have been developed in several languages (R, Excel, Java, Python, C#) to enable researchers to integrate RSML files into popular research workflow.

Keywords: root; format; image analysis; modeling

Introduction

By securing access to water and nutrients, root systems are generally recognized as having a critical influence on plant productivity (Lynch, 1995). As an example, in maize, historical increases in yield in the U.S. Corn Belt have been linked to an increase in root system size (Hammer et al., 2009). Tailoring root architecture is therefore thought to be a critical step towards dealing with extreme environmental conditions such as drought (Comas, 2013; Lobet et al., 2014) or nutrient-poor soils (Lynch, 2007; Postma and Lynch, 2011).

While precise root system architecture characterization methods have been studied in woody plant research for many years (Danjon et al., 1999; Danjon and Reubens, 2007; Danjon et al., 2013) physiological studies on smaller plants (e.g. *Arabidopsis thaliana*) have often neglected detailed root architecture, using mainly global estimators, such as total root length or the maximal depth of the root system convex hull (Galkovskyi et al., 2012). However, an increasing number of research questions now require precise quantification of root system architecture. As an example, nutrient or water deficiencies can have strong effects on root development (Al-Ghazi et al., 2003; Hammer et al., 2009; Péret et al., 2012; Gruber et al., 2013; Kellermeier et al., 2014) and only accurate root reconstruction allow the quantification of these effects. In addition, functional structural plant models are becoming increasingly popular to investigate the belowground ecophysiology of crops (Draye et al., 2010; Comas, 2013; Dunbabin et al., 2013; Lobet et al., 2014) and models require a precise quantification of root system architecture, either to evaluate root developmental parameters or as a direct model input.

Root systems architecture is generally described at three main levels (Godin and Sinoquet, 2005; Lynch, 2007; Postma and Lynch, 2011). The *geometry* of a root system describes the physical position, in space and time, of its component root axes. The *topology* describes the root system as a network and can be seen as the *backbone*, or *skeleton* of the root system. Finally, the successive *segments* which comprise individual root axes can be further characterized by their properties, such as the local root diameter, color, or the presence/absence of root hairs.

While these levels can be easily represented for simple root systems or single roots (fig. 1A.I), the complexity of the representation can dramatically increase for complex root systems and branched root axes (fig. 1A.II & III). In addition, while root system analysis is classically performed on 2D images of root projections, recent years have seen the development of 3D acquisition devices (Danjon et al., 1999; Danjon and Reubens, 2007; Iyer-Pascuzzi et al., 2010; Clark et al., 2011; Mooney et al., 2012; Danjon et al., 2013) which solve the issues of object occlusion, yet increase the complexity of the root system description (fig. 1B). Finally, the recourse to dynamic traits and the tracking of individual roots in root development studies requires elaborated time series data representation (fig. 1C).

The past few years have seen the development of a variety of solutions for the analysis of root system images (see Lobet et al. (2013) for an updated listing). Several of these solutions deal with root architecture *per se* and consider explicitly the morphological and topological properties of the root system (Table 1). Such a variety of software solutions reflects the coexistence of complementary approaches to the analysis of root systems. As a direct consequence of this diversity, many independent root system architecture representation and storage have been implemented, leading to multiple datasets lacking common structure, which restricts the possibilities to compare root system architecture structures or measurements obtained using different tools, or to validate new algorithms.

The Multi-Scale Tree Graph formalism (MTG, Godin and Caraglio (1998)) is a formalism used to represent the topology and the geometry of any type of plant architecture at different levels of organisation. This formalism has become a *de-facto* standard in the plant architecture community, encoding plant architecture and its development (Godin et al., 1999; Godin et al., 2005; Danjon and Reubens, 2007; Griffon and de Coligny, 2013) for a wide variety of plant architectures (e.g. root systems (Danjon et al., 1999), annual plants (Mündermann et al. , 2005; Fournier et al., 2010) and fruit trees (Guédon et al., 2001; Negrón and Contador, 2013)). In recent years, computational and mathematical models of growth, branching and architecture have been developed around this formalism on the basis of qualitative botanical knowledge (see Barthelemy and Caraglio (2007) for a review). The MTG is the central data-structure of the OpenAlea platform for FSPM (Fournier et al., 2010; Boudon et al., 2012; Garin et al., 2014) that eases communication between different models developed by different research groups. To achieve that level of genericity,

MTGs do not assume any specific type of plant architecture ontology, and can be adapted to each new protocol in a flexible manner. However, this flexibility may induce additional complexity where exchanging data between different groups of scientists using different protocols and software within a specific domain. For each new protocol for example, a modeler must define the number of scales, their meaning, the name of the attributes and how to encode them. While OpenAlea provides software solutions to manage this complexity, external software must implement and manage this complexity in their own programming environment and language.

Now, a new step must be taken to further improve the ability of researchers to acquire plant architecture data using different software, and to exchange and share this data. To achieve this, the research community needs to agree on a common biological language to build up shared databases and quantitative tools and to compare hypotheses and approaches. This is unfeasible at the level of genericity that was used in the design of MTGs. However, genericity can be achieved at the level of particular plants, plant parts (such as roots) or applications by developing specific ontologies on the top of MTGs.

This work introduces the Root System Markup Language (RSML), a unified language that enables root system architecture information storage based on the MTG formalism and XML standards. RSML aims to (i) accommodate the richness of root system architecture types and complexities (2D, 3D, time-series) and (ii) be open, cross platform and easy to implement in new tools and software.

At the time of writing, RSML support has been implemented into the following imaging or modelling suites: ArchiSimple (Pagès et al., 2014), OpenAlea (Pradal et al., 2008), RhizoBox (Leitner et al., 2010), RhizoScan (Diener et al., 2013), RootNav (Pound et al., 2013), RooTrak (Mairhofer et al., 2012; Mairhofer et al., 2013), RootSystemAnalyser (Leitner et al., 2013), R-SWMS (Javaux et al. 2008) and SmartRoot (Lobet et al., 2011).

Results and discussions

Description of the RSML format

The RSML format defines an xml file in which the topological, geometrical and numerical information describing a root system is stored.

In practice, the RSML format is split into metadata and the scene elements (Fig. 2). The metadata store experimental and technical information relevant to the scene, while the scene itself contains the root system representation. Supplementary data can be stored using properties (at the scene, plant or root level), functions along the root axes and annotations.

A brief outline of the main components of the RSML format is presented in the following sections. The RSML website (http://rootsystemml.github.io/) provides technical specifications of the format, as well as RSML examples and related image files which illustrate the possibilities offered by the RSML format. It also includes links to all software reading and writing RSML files. Supplemental data 1 contains a simple example of RSML file.

Metadata

The metadata specifies the experimental context in which an RSML file was constructed. It also provides a concise description of the file content, allowing documents to be quickly scanned, filtered or classified without reading the entirety of each file.

Included in the metadata element are a unique identifier for the scene, file information, and the real-world unit and resolution for the root conversion of pixel data. The <software> and <last-modified> elements allow tracking of changes when multiple software tools have handled the same document. <Property> and <function> definitions describe the associated properties or functions that will appear in the document. If the RSML file is one of many in a time series dataset, a <time-sequence> element identifies this set and the position of the file in the sequence.

Scene and topology

The scene represents a single image, possibly within a series of 2D or 3D images, and which will contains at least one root system. Within the scene, there is at least one plant, containing at least one root. Finally, roots may also contain additional child roots, i.e. lateral roots. RSML documents mimic this structure (fig. 2): the scene element will contain one or more plant elements, which in turn will contain one or more root elements. Further levels of root elements are used to show lateral roots of higher orders (fig. 2C). The topological aspects of the root system, such as the connexions between a primary root and its child lateral roots, are therefore represented through the nested structure of the scene element in the RSML document itself. All geometry and other measurements are stored as data attached to the relevant elements throughout the document (fig. 2C).

In order to ensure a uniform naming of the different root types and terms across different files, root type descriptors used in RSML refer to The Plant Ontology Database (Avraham et al. 2008), a plant ontology widely accepted within the plant community. The current root type terminology in the PO database is presented in table 2. The list is not exhaustive and might be extended with other terms from the PO database.

Root geometry

While the topology of the scene is implicit in the hierarchy of the document, geometric information is defined explicitly. Root elements contain a geometry element in which the root geometry is detailed as a polyline – a succession of linear segments. Scenes and plants contain no geometrical information; their geometries are the combined geometries of all child root elements.

RSML has been designed to allow sharing of root architecture between different software packages, where these systems may contain different representations of root geometry. We use the polyline as the primary geometric structure (each root element must include a polyline geometry). Each software that makes use of RSML is responsible for conversion between a polyline and any alternative structures that the software may use, such as continuous splines. Inside the geometrical description of a root, a polyline element contains an ordered list of points that provide the endpoints of the successive segments that make up the root. Each point element contains x and y attributes for 2D scenes, and an optional z attribute for 3D architectures. All geometric

units in RSML are given in pixel coordinates, referring directly to the image associated with this root system. RSML metadata can be used to provide the scaling necessary to convert into real-world units.

It may be the case that the conversion to polyline form comes at the expense of accuracy in a given software package. Should a certain application store geometry in spline form, for example, the conversion to a polyline will only approximate the curve. In this case additional geometries are permitted alongside, but not instead of, the polyline. These can take any form as long as they are contained within a single child of the geometry element. It should be noted that additional geometry types are included for the convenience of individual software developers. Other software that reads the RSML format need only examine the polyline form of the geometry, and may disregard any additional information. In this way, the portability of geometric information between software is ensured, but more specific structures are available if the RSML files are being used as a storage format for a particular application.

Root functions

It will often be desirable to attach additional information along the polyline. This would be the case of, e.g. the root diameter, the root age, the root hair length or the presence of nodules. In RSML, continuous functions are used to describe quantitative information as a function of the longitudinal position along the root axis.. The *function domain* is explicitly defined, and specifies the mapping between observed function values and their corresponding positions along the root (see supplemental figure 1 for a graphical example). Depending on the software implementation, the sample points of a function can be uniformly spread over a root length, or attached to a given position on the polyline, either using an index, or a length..

Through the use of functions, quantifiable variables can be added within an RSML document. Information that is not directly associated with a root geometry, or categorical information that cannot be provided as a function is instead stored in separate entities within the RSML specification that are described below.

Properties

Many aspects of root systems cannot be linked directly to root geometry, and are instead related to botanical entities. For example, while diameter is intrinsically linked to a position along the root, qualitative (long or short root, dead root) or global (length, insertion angle or position) information is better attached to the whole root as a

property. A <property> element contained within a scene, plant or root element specifies a measurement of that object. Properties can take one of numerous data types, allowing binary, integers, real types or text values. As properties may be specific to the generating software, the meaning of a property can be supplied within the document metadata. Properties might also be used for a more efficient parsing of the RSML document, enabling analysis tools to directly retrieve these pre-calculated properties without having to compute them from the topological and geometrical information.

Annotations

It may be useful to attach general information such as user observations to a given scene. For example, a region of a given scene could be marked as “out of focus”, letting software know that image analysis performed in this area may be less reliable. As properties, annotations are added as elements located within the corresponding scene, plant or root element. Each annotation includes a list of one or more points, representing the point, line or region of interest to which the annotation applies. A <value> element provides the text or numerical content of the annotation, and finally a <software> element specifies the software used to add that annotation.

Root Development

Root developmental processes (e.g. growth rate) are often analyzed using time-lapse image sequences. Preserving time series information is an important factor in many root phenotyping applications and requires maintaining an explicit link between successive images of the same plant. The RSML format allows images of a time-lapse sequence to be linked through the use of the <time-sequence> element in the metadata. The <index> element indicates its position in the time series.

Software that provides an explicit mapping between geometries in a time series can use the metadata can use to indicate this. Some software use such information to calculate change in parameters over time, e.g. elongation rate. Others (SmartRoot, RootSystemAnalyser) use previous time steps information to initialize subsequent reconstructions, improving root tracing efficiency by focusing on incremental information in the following image.

RSML Thesaurus

The RSML format does not impose a restricted set of properties or functions. However, a thesaurus has been defined to promote the use of standardized terms. Any supporting software is not required to process these terms. So, unlike the main RSML definition, new terms may be added to this thesaurus without changing the format itself. Addition to the thesaurus follows a traditional open-source protocol as described on the RSML website (http://rootsystemml.github.io).

Open-source packages for RSML analysis

Five open-source APIs have been created to read and parse RSML datafiles from within C#, Excel (fig. 3A), R (fig. 3B), ImageJ (fig. 3C) and Python (fig. 3D). The aim of these packages is not to carry the analysis of the root system data, but to provide end-users with commonly used data structures within popular data analysis pipelines. These packages have been released as open-source to allow users to adapt them to their needs. They are available through the RSML website.

RSML enables common pipelines for root systems analysis and modeling

The RSML format provides plant researchers with a central paradigm connecting image analysis tools, data analysis pipelines and modeling platforms (fig. 4). Data generated by RSML-compliant tools can be reused in others, facilitating data transfer between researchers and groups. We provide here three examples where RSML is used to interface different analysis pipelines.

In the first example, RSML was used to transfer root architecture information between root image analysis tools (fig. 5). A root image, containing several plants (fig. 5A), was traced using RootNav (fig. 5B). RootNav features an efficient root tracing algorithm, but does not compute a measure of root diameter along each root, a measurement that might be required in some experiment. The RSML file generated by RootNav was imported in SmartRoot, which automatically computes diameter measurements upon loading an RSML file that does not contain that information (fig. 5C). The resulting RSML file was imported into the R statistical computing environment (R Core Team) for analysis. The profile of lateral root length along the primary root axis (fig. 5D) was computed using the tracing originally performed by RootNav. The primary and lateral root diameters distributions (fig. 5E) were computed using the data computed by SmartRoot. This example illustrates the complementarity

of existent root image analysis tools and how the RSML format enables this complementarity to be exploited by researchers.

Today's science faces an increasing demand for reproducibility and standardized analysis pipelines. We believe the existence of a standard format for root architecture will enable easier reproducibility practice amongst researchers and allow the comparison of multiple datasets, even those coming from different sources. In the second example, the image shown in figure 6A was analyzed using RhizoScan, RootSystemAnalyser, RootNav and SmartRoot. The RSML files generated by the different tools were exported into a single datafile and analyzed using R (R Core Team). Supplemental figure 2 shows the comparison of the measured primary root length, lateral root length, lateral insertion angles and lateral insertion positions for each software. This example illustrates how a shared format can streamline the validation of new algorithms and the creation of benchmark datasets with which to validate them.

In the third example, we illustrate the use of RSML for data storage and sharing between modeling tools. Figure 6A shows the visual output of an *Anagallis femina* root system simulated by RootBox (Leitner et al., 2010). The simulated root architecture was stored as an RSML data file, and converted into the MTG data-structure (Godin and Caraglio, 1998) in the OpenAlea platform (Pradal et al., 2008). Taking advantage of the geometric modules of PlantGL (Pradal et al., 2009), the 2D or 3D convex hull of the root system can easily be calculated (fig. 6B). The same RSML file was used in R-SWMS (Javaux et al. 2008) to simulate water flow in the soil-root system, hence allowing the testing of the functional performance (in this case water uptake) of the simulated root system.

These three examples highlight the potential role of the RSML format as a cornerstone in analysis pipelines and show how it can hasten data (both simulated and experimental) exchange between researchers.

RSML promotes the use of a central repository for root architecture data

Many experiments on root architecture, and so a great number of software tools developed to analyze them, focus on the limited number of root architecture

parameters that can be calculated without an explicit root model. To extract other parameters of interest (e.g. lateral root length), a complete tracing of the root system is often required, including a hierarchical model of the root structure. However, the tracing of a complete root system can be time consuming. It is therefore highly beneficial to reuse previous root datasets in the quantification of other traits demanded in different experiments. Due to the lack of compatibility between many historic datasets, this re-analysis is often only possible by re-construcing the complete dataset, which is at best time-consuming and, at worst, impossible. By storing root architecture in a common format, desired root traits can be calculated quickly over large datasets captured with a variety of software, regardless of the traits that were considered when that dataset was first analyzed.

We believe that the adoption of RSML will encourage the creation of central repositories for root architecture data, similar to those that exist in other domains. In molecular biology, it is even mandatory to upload gene expression datasets to a database such as EBI-ArrayExpress (Rustici et al., 2012) prior to publication. Those publicly accessible repositories are frequently queried by the scientific community. The development of similar public database for costly and valuable root architecture data would increase the pace and efficiency of root research.

This central repository can also be used as a benchmark to compare and evaluate computer programs used to reconstruct architectural data. New algorithms and software could be assessed to ensure that the datasets produced are scientifically valid. This benchmark will also greatly accelerate the impact and the adoption of new, independently developed algorithms for automatic digital reconstruction of root system architecture. In neuroscience, the DIADEM Challenge (DIgital reconstructions of Axonal and DEndritic Morphology) addresses a similar need (Parekh and Ascoli, 2013).

Using RSML format to store root system data without architectural information

The RSML format was initially defined an efficient storage mechanism for detailed root system representations. As such, the explicit topology and geometry of the root system can be encoded. However, it is important to note that the RSML format can also be used for the storage of root system data that does not contain geometrical

and/or topological information. As an example, the recently developed DIRT image analysis toolbox (Bucksch et al. 2014) is able to extract multiple metrics from images of field-grown crown roots. The RSML format could still be used in this case, by taking advantage of its underlying multiscale formalism, i.e. the fact that properties may be set at any level of detail (scene, plant, root).

Conclusions

The Root System Markup Language (RSML) presented in this paper facilitates the sharing of root architectures between software, experiments and research groups. RSML accommodates for a wide array of root architecture complexity (ranging from 2D projections of single roots through to 3D representations of complete root systems), at varying levels of detail. The RSML format stores root topology (parent-child relationships), morphological properties (positions in space and time, length), and virtually any type of additional information used to describe root segments (e.g. diameter, color, age) separately, linking them to form a coherent representation.

The RSML format is currently implemented within five root image analysis software (RhizoScan, RootNav, RooTrak, RootSystemAnalyser and SmartRoot), three functional-structural root models (RootBox, ArchiSimple and R-SWMS), and is based on the MTG format used in the OpenAlea platform. Five open-source packages were developed for the analysis and visualization of RSML data files (in C#, R, Excel, Java and Python). We believe the availability of RSML will encourage the creation of common analysis pipelines for root architecture information, enabling better data sharing between root researchers, and will facilitate the creation of shared database of root architecture information.

A complete description of the Root System Markup Language, along with examples and API packages, are available at http://rootsystemml.github.io

Literature cited

Al-Ghazi Y, Muller B, Pinloche S, Tranbarger TJ, Nacry P, Rossignol M, Tardieu F, Doumas P (2003) Temporal responses of Arabidopsis root architecture to phosphate starvation: evidence for the involvement of auxin signalling. Plant, Cell and Environment **26**: 1053–1066

Armengaud P, Zambaux P, Hills A, Sulpice R, Pattison RJ, Blatt MR, Amtmann A (2009) EZ-Rhizo: integrated software for the fast and accurate measurement of root system architecture. The Plant Journal **57**: 945–956

Avraham, S., Tung, C.-W., Ilic, K., Jaiswal, P., Kellogg, E.A., Mccouch, S., Pujar, A., Reiser, L., Rhee, S.Y., Sachs, M.M., Schaeffer, M., Stein, L., Stevens, P., Vincent, L., Zapata, F., Ware, D. (2008) The Plant Ontology Database: A community resource for plant structure and developmental stages controlled vocabulary and annotations. Nucleic Acids Research, 36 (SUPPL. 1): D449-D454.

Barthelemy D, Caraglio Y (2007) Plant Architecture: A Dynamic, Multilevel and Comprehensive Approach to Plant Form, Structure and Ontogeny. Annals of Botany **99**: 375–407

Boudon F, Pradal C, Cokelaer T, Prusinkiewicz P, Godin C (2012) L-py: an L-system simulation framework for modeling plant architecture development based on a dynamic language. Frontiers in Plant Science **3**: 1–20

Bucksch A, Burridge J, York LM, Das A, Nord E, Weitz JS, Lynch JP (2014). Image-based high-throughput field phenotyping of crop roots. Plant Physiology. **166:** 470-486

Clark RT, Famoso AN, Zhao K, Shaff JE, Craft EJ, Bustamante CD, McCouch SR, Aneshansley DJ, Kochian LV (2012) High-throughput 2D root system phenotyping platform facilitates genetic analysis of root growth and development. Plant, Cell and Environment **36**: 454–466

Clark RT, MacCurdy R, Jung J, Shaff J, McCouch SR, Aneshansley D, Kochian L (2011) 3-Dimensional Root Phenotyping with a Novel Imaging and Software Platform. Plant Physiology **156**: 455–465

Comas LH (2013) Root traits contributing to plant productivity under drought. Frontiers in Plant Science **4**: 1–16

Danjon F, Khuder H, Stokes A (2013) Deep Phenotyping of Coarse Root Architecture in R. pseudoacacia Reveals That Tree Root System Plasticity Is Confined within Its Architectural Model. PLoS ONE **8**: e83548

Danjon F, Reubens B (2007) Assessing and analyzing 3D architecture of woody root systems, a review of methods and applications in tree and soil stability, resource acquisition and allocation. Plant and Soil **303**: 1–34

Danjon F, Sinoquet H, Godin C, Colin F, Drexhage M (1999) Characterisation of structural tree root architecture using 3D digitising and AMAPmod software. Plant

and Soil **211**: 241–258

Diener J, Nacry P, Périn C, Dievart A, Draye X, Boudon F, Gaujon A, Godin C (2013) An automated image-processing pipelinefor high-throughput analysis of root architecture in OpenAlea. *In* 7th International Conference on Functional-Structural Plant Models. pp 85–87

Draye X, Kim Y, Lobet G, Javaux M (2010) Model-assisted integration of physiological and environmental constraints affecting the dynamic and spatial patterns of root water uptake from soils. Journal of Experimental Botany **61**: 2145–2155

Dunbabin VM, Postma JA, Schnepf A, Pagès L, Javaux M, Wu L, Leitner D, Chen YL, Rengel Z, Diggle AJ (2013) Modelling root–soil interactions using three–dimensional models of root growth, architecture and function. Plant and Soil **372**: 93–124

Fournier C, Pradal C, Louarn G, Combes D, Soulié J-C, Luquet D, Boudon F, Chelle M (2010) Building modular FSPM under OpenAlea: concepts and applications. 6th International Workshop on Functional-Structural Plant Models 109–112

French AP, Ubeda-Tomas S, Holman T, Bennett M, Pridmore T (2009) High-Throughput Quantification of Root Growth Using a Novel Image-Analysis Tool. Plant Physiology **150**: 1784–1795

Galkovskyi T, Mileyko Y, Bucksch A, Moore B, Symonova O, Price CA, Topp CN, Iyer-Pascuzzi AS, Zurek PR, Fang S, et al (2012) GiA Roots: software for the high throughput analysis of plant root system architecture. BMC Plant Biology **12**: 116

Garin G, Fournier C, Andrieu B, Houlès V, Robert C, Pradal C (2014) A modelling framework to simulate foliar fungal epidemics using functional–structural plant models. *Annals of botany* **114**: 795–812.

Godin C, Caraglio Y (1998) A Multiscale Model of Plant Topological Structures. Journal of theoretical Biology **191**: 1–46

Godin C, Costes E, Sinoquet H (1999). A method for describing plant architecture which integrates topology and geometry. Annals of Botany, 84(3), 343–357.

Godin C, Costes E, Sinoquet H (2005). Plant architecture modelling. In C. G. N. Turnbull, Plant architecture and its manipulation. Wiley-Blackwell.

Godin C, Sinoquet H (2005) Functional-structural plant modelling. New Phytologist **166**: 705–708

Griffon S, de Coligny F (2013) AMAPstudio: An editing and simulation software suite for plants architecture modelling. Ecological Modelling 1–8

Gruber BD, Giehl RFH, Friedel S, Wiren von N (2013) Plasticity of the Arabidopsis root system under nutrient deficiencies. Plant Physiology **163**: 161–179

Guédon Y, Barthelemy D, Caraglio Y, Costes E (2001) Pattern analysis in branching and axillary flowering sequences. Journal of theoretical Biology **212**:

481–520

Hammer G, Dong Z, Mclean G, Doherty A, Messina C, Schussler J, Zinselmeier C, Paszkiewicz S, Cooper M (2009) Can Changes in Canopy and/or Root System Architecture Explain Historical Maize Yield Trends in the U.S. Corn Belt? Crop Science **49**: 299–312

Iyer-Pascuzzi A, Symonova O, Mileyko Y, Hao Y, Belcher H, Harer J, Weitz J, Benfey PN (2010) Imaging and Analysis Platform for Automatic Phenotyping and Trait Ranking of Plant Root Systems. Plant Physiology **152**: 1148–1157

Javaux M, Schroeder T, Vanderborght J, Vereecken H (2008). Use of a three-dimensional detailed modeling approach for predicting root water uptake. Vadose Zone Journal, **7**, 1079–1088

Kellermeier F, Armengaud P, Seditas TJ, Danku J, Salt DE, Amtmann A (2014) Analysis of the Root System Architecture of Arabidopsis Provides a Quantitative Readout of Crosstalk between Nutritional Signals. The Plant Cell **26**: 1480–1496

Le Bot J, Serra V, Fabre J, Draye X, Adamowicz S, Pagès L (2010) DART: a software to analyse root system architecture and development from captured images. Plant and Soil **326**: 261–273

Leitner D, Felderer B, Vontobel P, Schnepf A (2013) Recovering root system traits using image analysis - Exemplified by 2-dimensional neutron radiography images of lupine. Plant Physiology **164**: 24–35

Leitner D, Klepsch S, Bodner G, Schnepf A (2010) A dynamic root system growth model based on L-Systems. Plant and Soil **332**: 117–192

Lobet G, Couvreur V, Meunier F, Javaux M, Draye X (2014) Plant Water Uptake in Drying Soils. Plant Physiology **164**: 1619–1627

Lobet G, Draye X, Périlleux C (2013) An online database for plant image analysis software tools. Plant Methods **9**: 1–7

Lobet G, Pagès L, Draye X (2011) A Novel Image Analysis Toolbox Enabling Quantitative Analysis of Root System Architecture. Plant Physiology **157**: 29–39

Lynch JP (1995) Root Architecture and Plant Productivity. Plant Physiology **109**: 7

Lynch JP (2007) Roots of the second green revolution. Australian Journal of Botany **55**: 493–512

Mairhofer S, Zappala S, Tracy S, Sturrock C, Bennett MJ, Mooney SJ, Pridmore TP (2013) Recovering complete plant root system architectures from soil via X-ray mu-Computed Tomography. Plant Methods **9**: 1–7

Mairhofer S, Zappala S, Tracy SR, Sturrock C, Bennett M, Mooney SJ, Pridomore T (2012) RooTrak: Automated Recovery of Three-Dimensional Plant Root Architecture in Soil from X-Ray Microcomputed Tomography Images Using Visual Tracking. Plant Physiology **158**: 561–569

Mooney SJ, Pridmore TP, Helliwell J, Bennett MJ (2012) Developing X-ray Computed Tomography to non-invasively image 3-D root systems architecture in

soil. Plant and Soil **352**: 1–22

Mundermann L, Erasmus Y, Lane B, Coen E, Prusinkiewicz P (2005). Quantitative modeling of Arabidopsis development. Plant Physiology, 139(2), 960–968

Negrón C, Contador L (2013) Systematic Analysis of Branching Patterns of Three Almond Cultivars with Different Tree Architectures. Journal of the American Society for Horticultural Science **138**: 407–415

Pace J, Lee N, Naik HS, Ganapathysubramanian B, Lübberstedt T (2014). Analysis of Maize (Zea mays L.) Seedling Roots with the High-Throughput Image Analysis Tool ARIA (Automatic Root Image Analysis). PLoS ONE **9** : e108255

Pagès L, Bécel C, Boukcim H, Moreau D, Nguyen C, Voisin A-S (2014). Calibration and evaluation of ArchiSimple, a simple model of root system architecture. Ecological Modelling **290**: 76–84.

Parekh R, Ascoli GA (2013) Neuronal Morphology Goes Digital: A Research Hub for Cellular and System Neuroscience. Neuron **77**: 1017–1038

Péret B, Li G, Zhao J, Band LR, Voss U, Postaire O, Luu D-T, Da Ines O, Casimiro I, Lucas M, et al (2012) Auxin regulates aquaporin function to facilitate lateral root emergence. Nat Cell Biol **14**: 991–998

Postma JA, Lynch JP (2011) Root Cortical Aerenchyma Enhances the Growth of Maize on Soils with Suboptimal Availability of Nitrogen, Phosphorus, and Potassium. Plant Physiology **156**: 1190–1201

Pound MP, French AP, Atkinson J, Wells DM, Bennet JM, Pridmore T (2013) RootNav: Navigating images of complex root architectures. Plant Physiology **162**: 1802–1814

Pradal C, Dufour-Kowalski S, Boudon F, Fournier C, Godin C (2008) OpenAlea: a visual programming and component-based software platform for plant modelling. Functional Plant Biol **35**: 751–760

Pradal C, F B, Nouguier C, Chopard J, Godin C (2009) PlantGL: A Python-based geometric library for 3D plant modelling at different scales. Graphical Models **71**: 1–21

R Core Team R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria

Rustici G, Kolesnikov N, Brandizi M, Burdett T, Dylag M, Emam I, Farne A, Hastings E, Ison J, Keays M, et al (2012) ArrayExpress update--trends in database growth and links to data analysis tools. Nucleic Acids Research **41**: D987–D990

Silva D, F B, Godin C, Puech O, Smith C, Sinoquet H (2006) A Critical Appraisal of the Box Counting Method to Assess the Fractal Dimension of Tree Crowns. *In* G Bebis, R Boyle, B Parvin, D Koracin, P Remagnino, A Nefian, G Meenakshisundaram, V Pascucci, J Zara, J Molineros, et al., eds, Lecture Notes in Computer Science. Springer Berlin Heidelberg, pp 751–760

The Plant Ontology Consortium (2002) The Plant Ontology™ Consortium and Plant

Ontologies. Comparative and Functional Genomics **3**: 137–142

Tables

Table 1: Description of existing root system architecture image analysis tools

Software	Automation	Image type	Storage	Topology	Diameter	Time series	RSML support	Reference
ARIA	automated	2D	XLS	**yes**	no	no	no	(Pace et al . 2014)
EZ-Rhizo	automated	2D	SQL	**yes**	no	no	no	(Armengaud et al., 2009)
DART	manual	2D	TXT	**yes**	no	no	no	(Le Bot et al., 2010)
OpenAlea.RhizoScan	automated	2D	MTG	**yes**	no	**yes**	**yes**	(Diener et al., 2013)
RootNav	semi-automated	2D	-	**yes**	no	no	**yes**	(Pound et al., 2013)
RootReader2D	automated	2D	XML	**yes**	no	no	no	(Clark et al., 2012)
RootReader3D	automated	3D	XML	**yes**	no	no	no	(Clark et al., 2011)
RootSystemAnalyser	automated	2D	MAT	**yes**	**yes**	**yes**	**yes**	(Leitner et al., 2013)
RooTrak	automated	3D	-	**yes**	no	no	**yes**	(Mairhofer et al., 2012)
RootTrace	automated	2D	-	**yes**	no	**yes**	in progress	(French et al., 2009)
SmartRoot	semi-automated	2D	XML	**yes**	**yes**	**yes**	**yes**	(Lobet et al., 2011)

Table 2: Plant Ontology terms currently used into the RSML format. This list is not exhaustive as any term contained into the Plant Ontology database (www.plantontology.org) could be used.

Plant Ontology name	Plant Ontology accession
root	PO:0009005
basal root	PO:0025002
embryo root	PO:0000045
lateral root	PO:0020121
primary root	PO:0020127
shoot-borne root	PO:0000042
tuberous root	PO:0025523

Figures

Figure 1. Root system architecture description. A. Single root image: I. single root, II. single root axis and III. root system (e.g. for monocots). B. 3D image stack. C. Time series. D. Example of parameters used to described root architecture parameters. First order roots are shown in red, while second order roots in green.

Figure 2: visual representation of the RSML structure. A. Original image B. Graphical representation of the structure. Topology (primary root in red and lateral roots in green), geometry and properties are represented at different levels. C. Schematic representation of an RSML file structure. D. Representation of the coupling between the root geometry and its associated functions (here the diameter). Dotted lines represent data from the same point in a polyline.

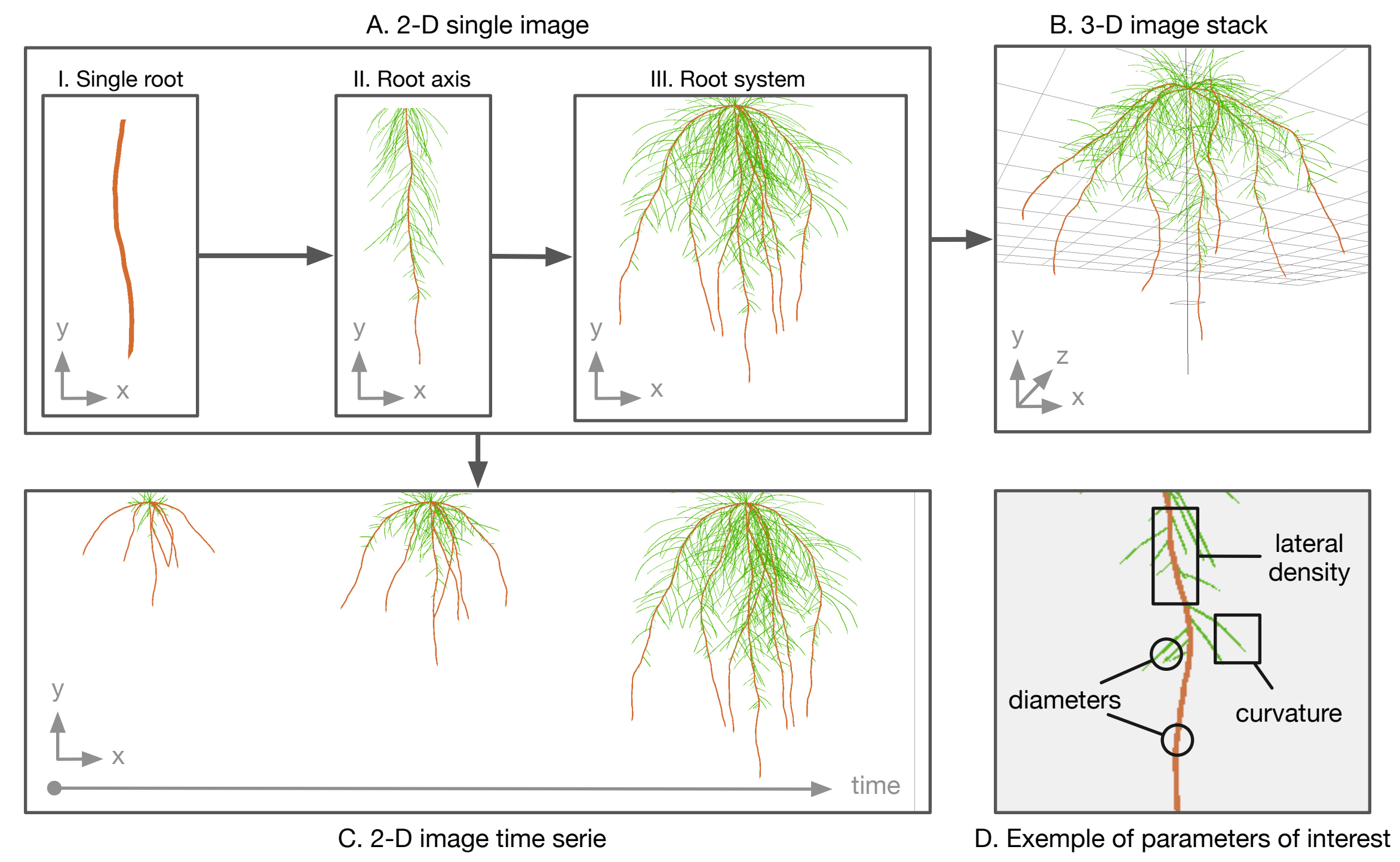
Figure 3: Visual output of different RSML analysis package. A. Excel plugin. B. R package. C. ImageJ plugin, D.Python package.

Figure 4: Analysis pipeline enabled by the RSML format. Dotted arrows represent connections that are not yet implemented.

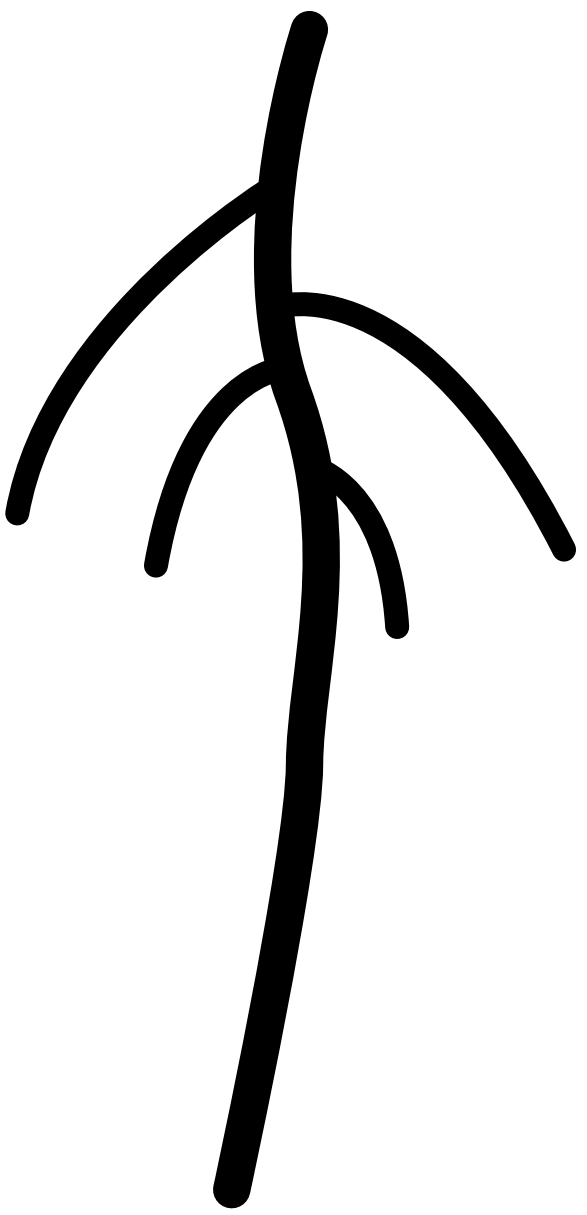
Figure 5: Example workflow enabled by the RSML format. A. Original image of Arabidopsis plants grown in petri dish. B. Screenshot of a root tracing done using RootNav. C. The RSML generated by RootNav was opened using SmartRoot, that computed the root diameter (which is not calculated by RootNav). D. R-generated graph showing lateral root length depending on the insertion position from the primary root base. These data were computed by RootNav. Dashed line represents the moving average across the dataset. E. R-generated histograms comparing the diameter of the primary and lateral roots. This data was computed by SmartRoot.

Figure 6: Using the RSML format to share data between modeling tools. A. Root system of *Anagallis femina,* simulated using RootBox (Leitner et al., 2010). B. The RSML file exported by RootBox was opened in PlantGL (Pradal et al., 2009) to compute the convex hull of the root

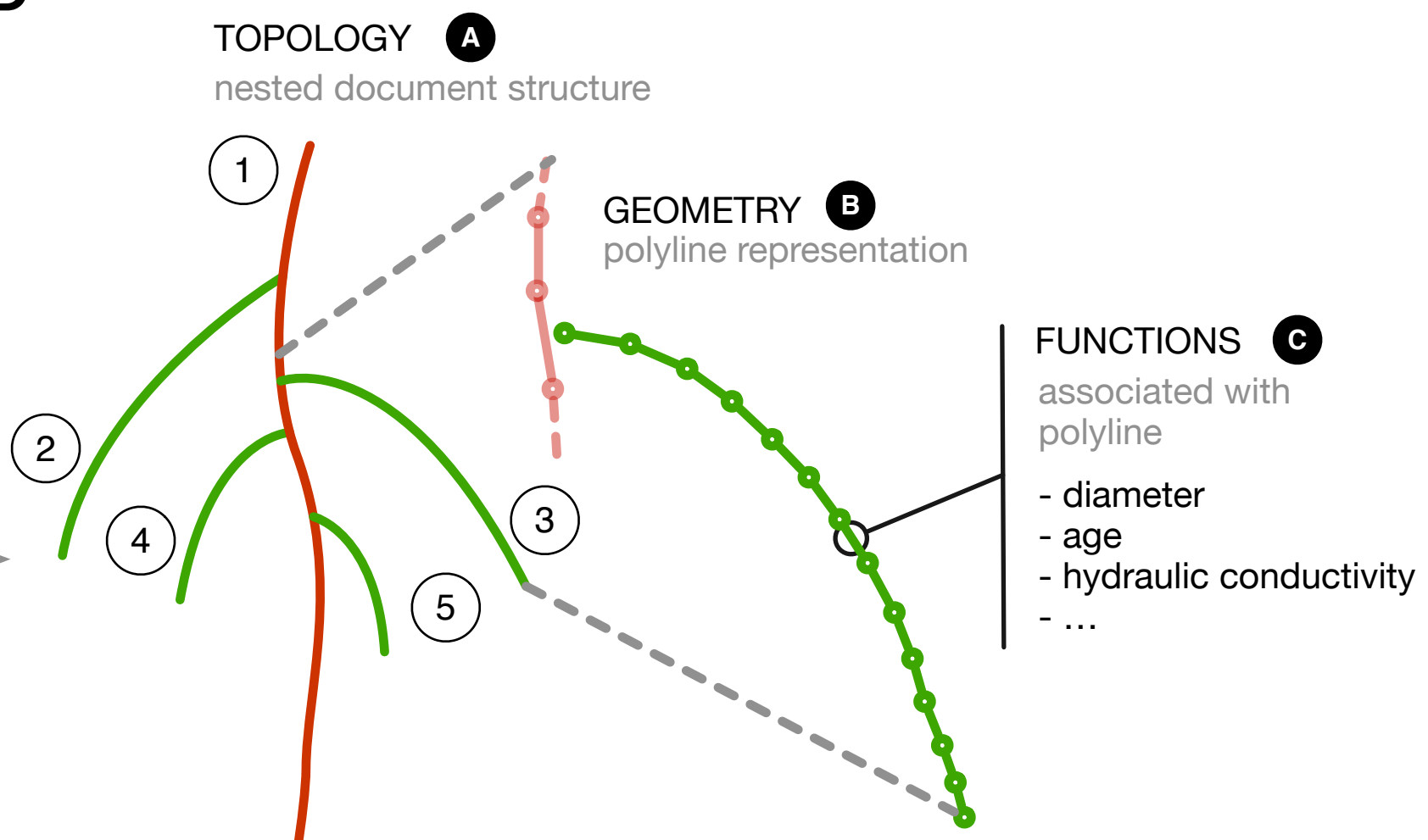
system (yellow lines). C.The same RSML file was used by R-SWMS (Javaux et al. 2008) to simulate soil water depletion in 24 hours.


A. 2-D single image
I. Single root
II. Root axis
III. Root system
y
x
B. 3-D image stack
z
time
C. 2-D image time serie
lateral density
diameters
curvature
D. Exemple of parameters of interest

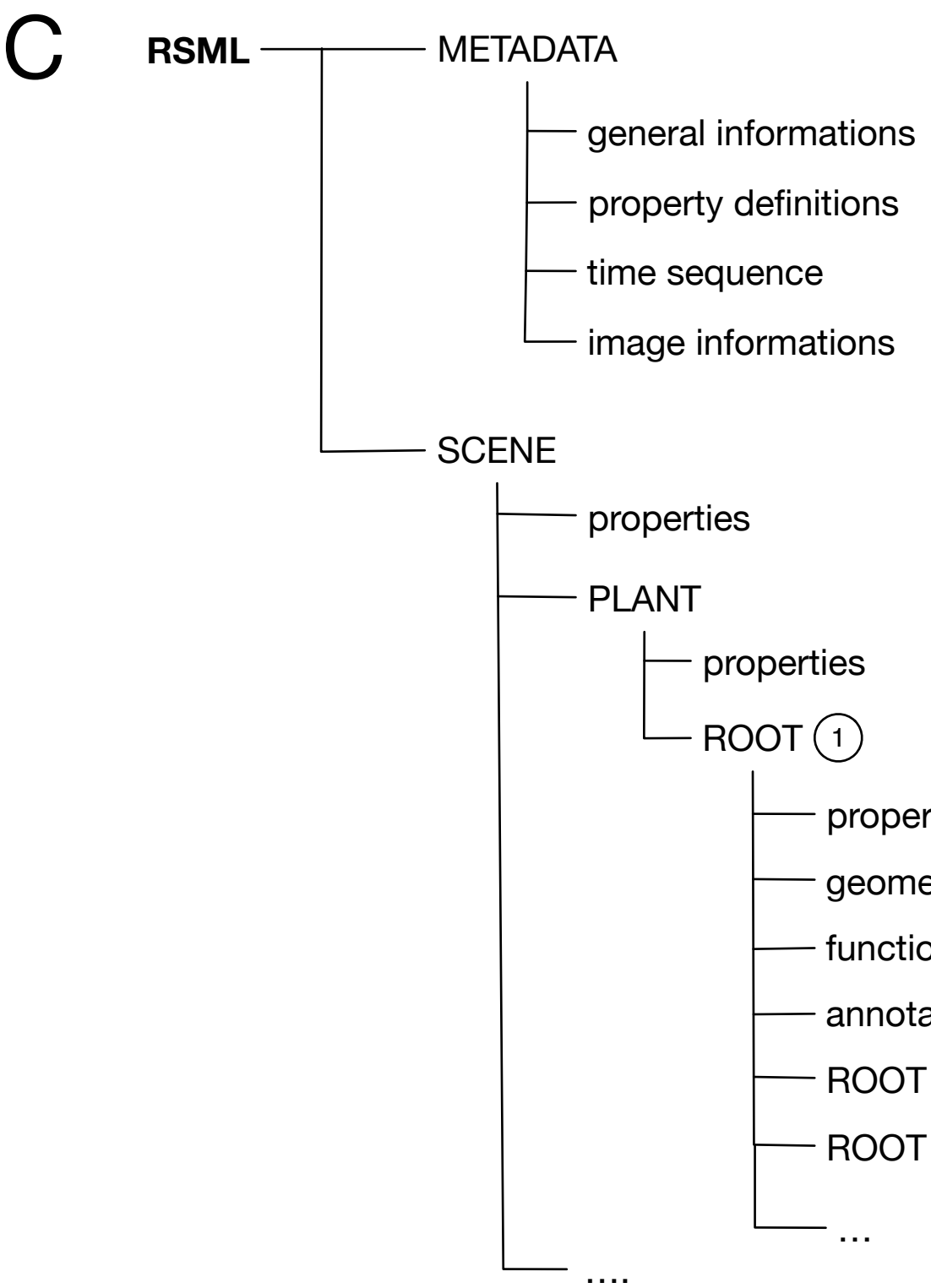
A



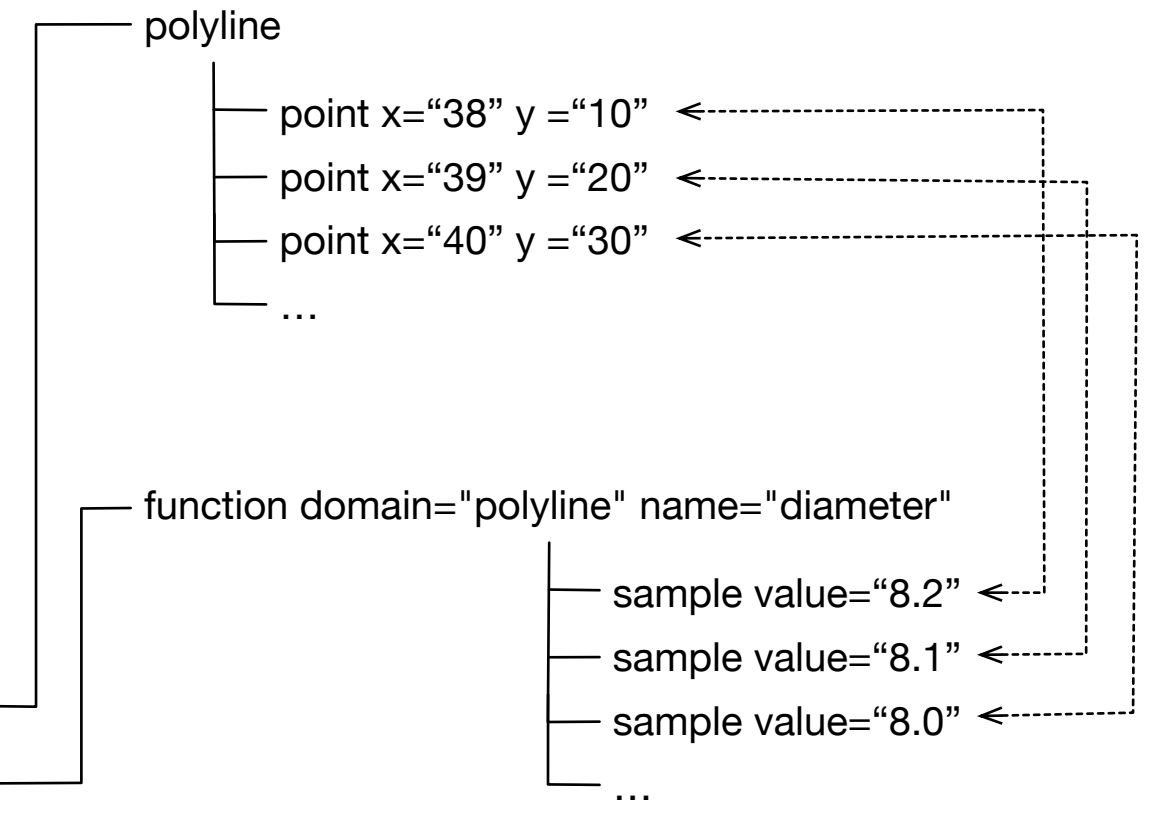
B


TOPOLOGY A
nested document structure
1
2
3
4
5
GEOMETRY B
polyline representation
FUNCTIONS C
associated with polyline
- diameter
- age
- hydraulic conductivity
- ...

C

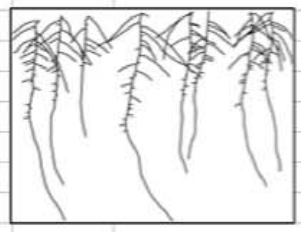

RSML
METADATA
general informations
property definitions
time sequence
image informations
SCENE
properties
PLANT
properties
ROOT 1
properties
geometry B
functions C
annotations
ROOT 2 A
ROOT 3
...
....

D


polyline
point x="38" y ="10"
point x="39" y ="20"
point x="40" y ="30"
...
function domain="polyline" name="diameter"
sample value="8.2"
sample value="8.1"
sample value="8.0"
...

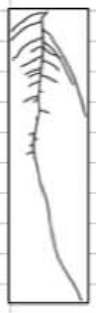
A

RSML Metadata	
Version	1
Unit	pixel
Resolution	1
Last Modified	18/07/2014
Software	RootNav
User	stxrt4



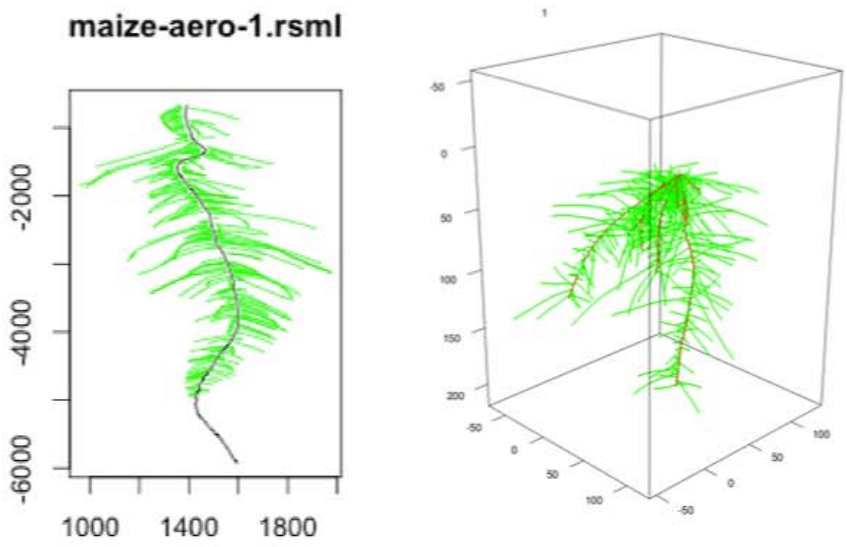
Plant 1	Primary	Lateral	Length	Emergence Angle	Tip Angle	Tortuosity
	1		1293,38	-39,15	22,19	1,03
		1	165,14	-86,97	-44,13	1,04
		2	14,32	77,91	77,91	1
		3	8,06	-97,13	-97,13	1
		4	28,61	-71,57	-71,57	1,01
		5	158,85	-69,93	-43,03	1,03
		6	54,49	-73,3	-60,85	1,01
		7	83,56	69,82	51,02	1,01
		8	61,64	-71,94	-65,23	1
		9	22,58	-77,2	-77,2	1
		10	424,05	54,58	15,58	1,02
		11	99,33	-64,07	-62,12	1

Plant Summary	
Total Length	3240,37
Primary Root Count	1
Lateral Root Count	22
Average Primary Length	1293,38
Average Lateral Length	88,5
Tortuosity	1,01
Convex Hull Area	226065,5



B

0.11871863	58.4786317	8.361149	-1.00000	-1.000000
0.16590367	59.4780386	8.140781	-1.00000	-1.000000
0.59281250	76.2882953	7.846747	-1.00000	-1.000000
0.09347127	64.4393241	7.586923	-1.00000	-1.000000
0.08775745	67.6803855	7.329572	-1.00000	-1.000000
0.33014384	76.2529286	7.329572	-1.00000	-1.000000
1.18519580	72.1459085	7.115103	-1.00000	-1.000000
0.78909403	35.6579861	6.875132	-1.00000	-1.000000
0.86944157	73.4802269	6.588412	-1.00000	-1.000000
0.53581080	52.3275272	6.588412	-1.00000	-1.000000
0.78632150	64.3878381	6.265322	-1.00000	-1.000000
0.47494054	90.4249154	6.588412	-1.00000	-1.000000
0.34999987	81.8601724	6.265322	-1.00000	-1.000000
0.58094580	62.2500845	5.986844	-1.00000	-1.000000
1.63342450	72.9434014	5.986844	-1.00000	-1.000000



C

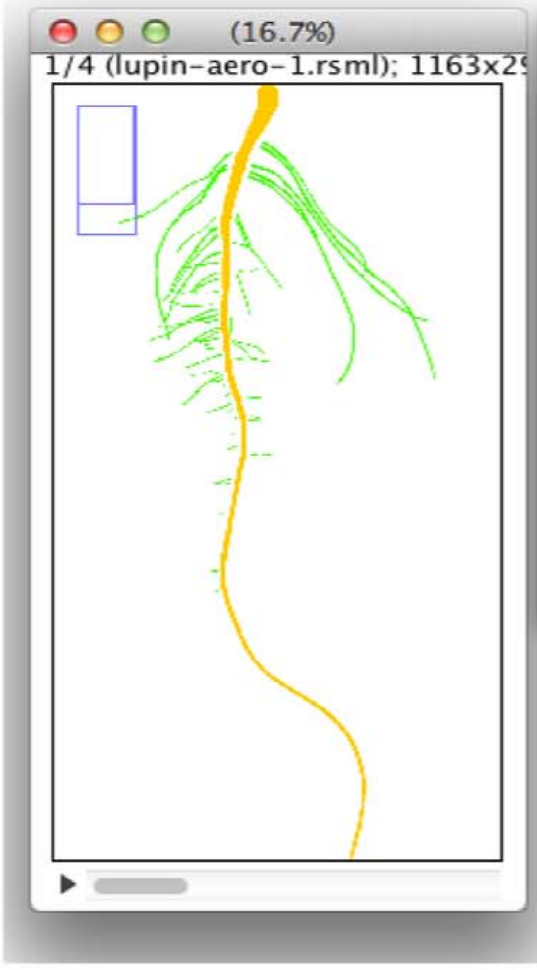


Image data

image	tot_root_length	n_primary	tot_p

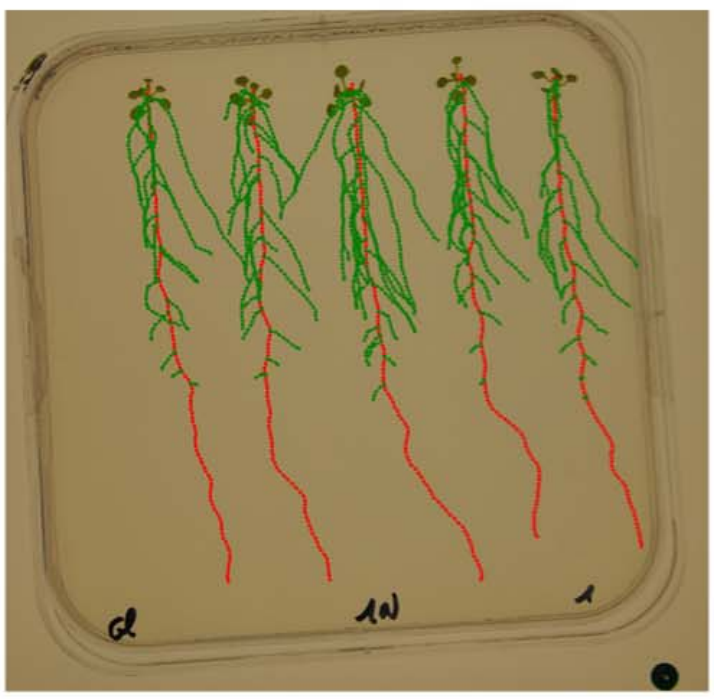
Root data

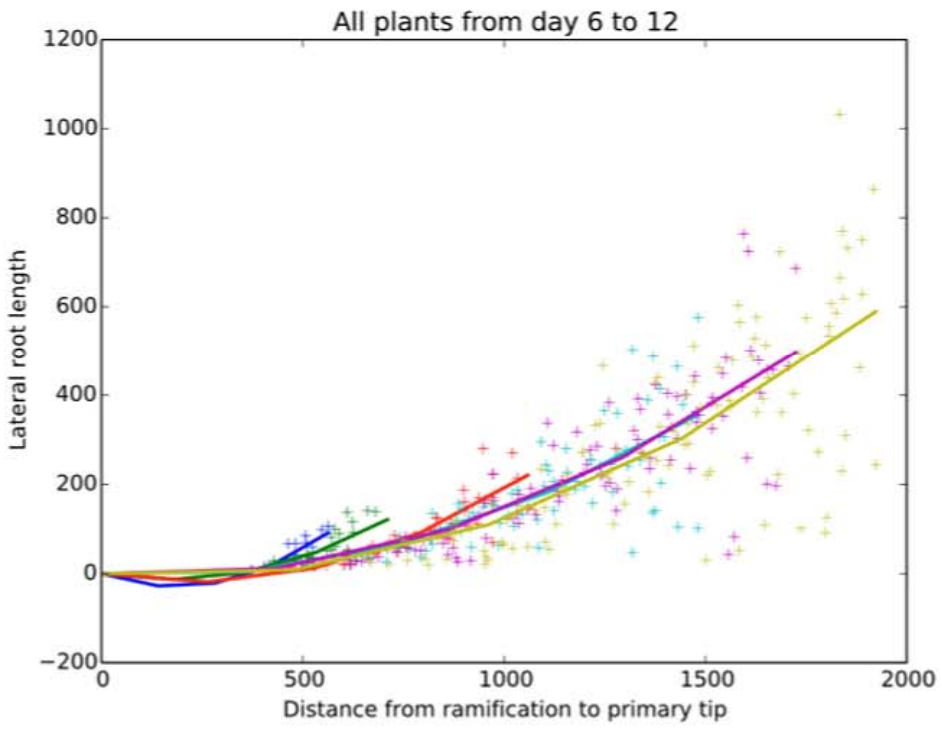
urface	volume	diameter	root_order	root_ontology	par

Node data

x	y	theta	diameter	distance_from_base	distanc
4.699	0.182	4.700	0.334	0.000	26.557
4.695	0.509	4.351	0.448	0.326	26.231
4.555	0.880	4.254	0.389	0.723	25.834
4.443	1.107	4.274	0.337	0.976	25.581
4.367	1.270	4.291	0.324	1.156	25.401
4.143	1.769	4.482	0.253	1.703	24.855
4.067	2.093	4.452	0.279	2.036	24.522
3.923	2.633	4.505	0.241	2.595	23.962
3.824	3.104	4.541	0.224	3.076	23.481
3.765	3.447	4.703	0.216	3.424	23.133
3.763	3.663	4.759	0.190	3.640	22.917
3.781	4.041	4.698	0.165	4.018	22.539
3.776	4.363	4.696	0.160	4.341	22.216
3.771	4.694	4.726	0.176	4.672	21.885
3.776	5.057	4.646	0.177	5.035	21.523
3.753	5.403	4.687	0.151	5.381	21.176

D





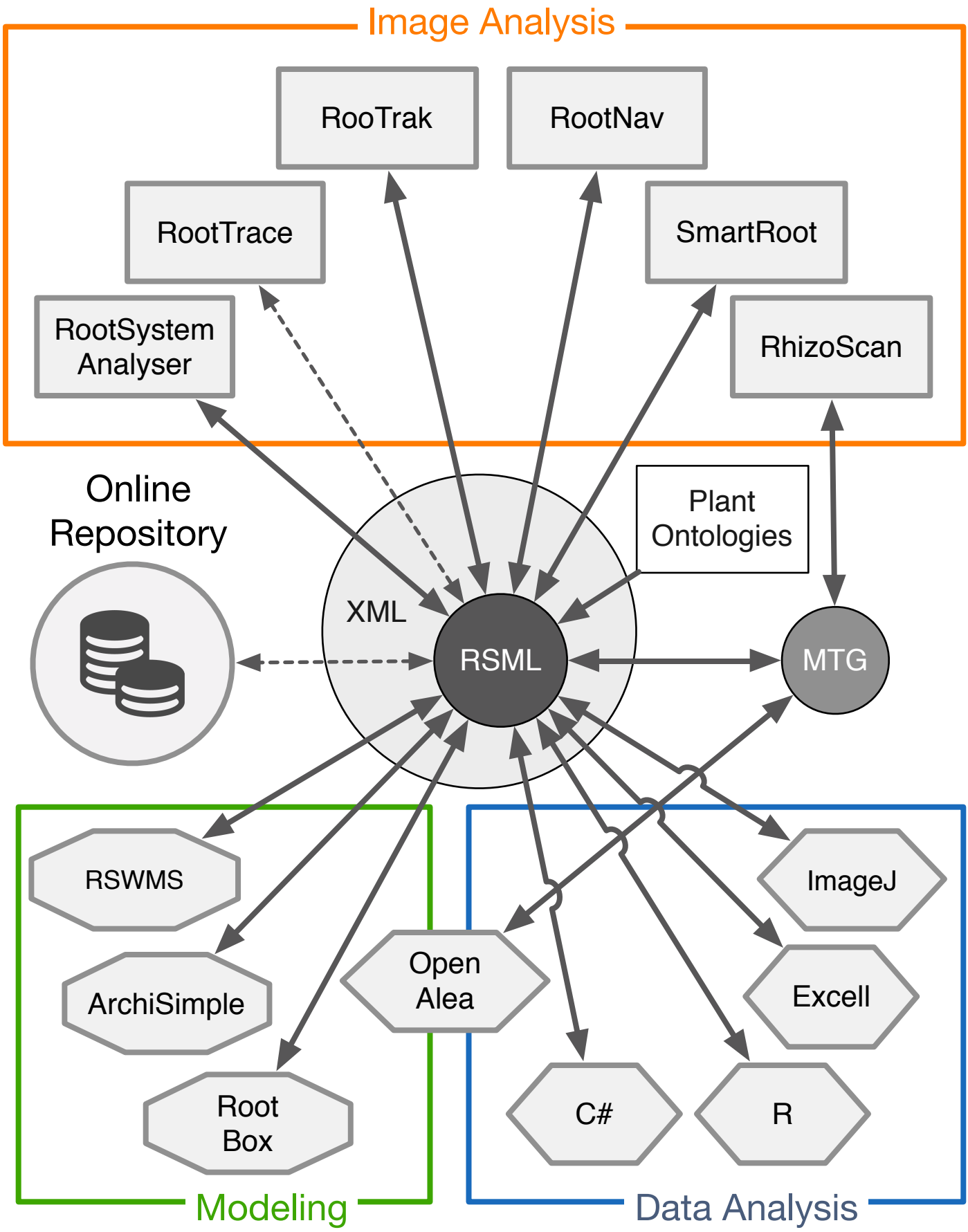
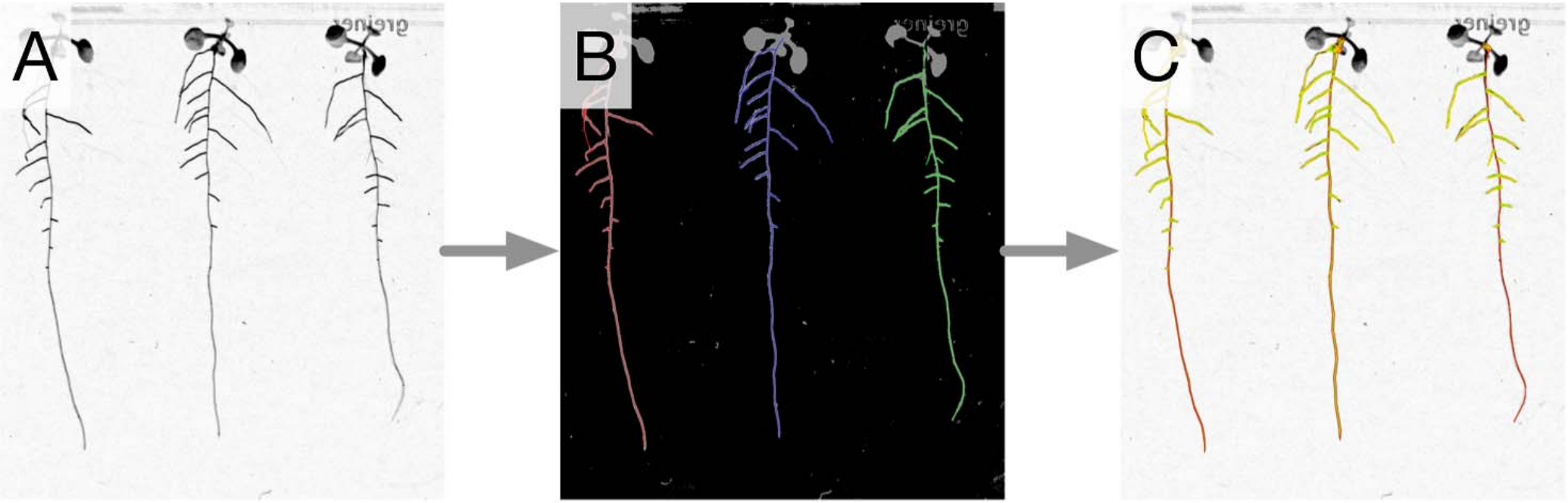
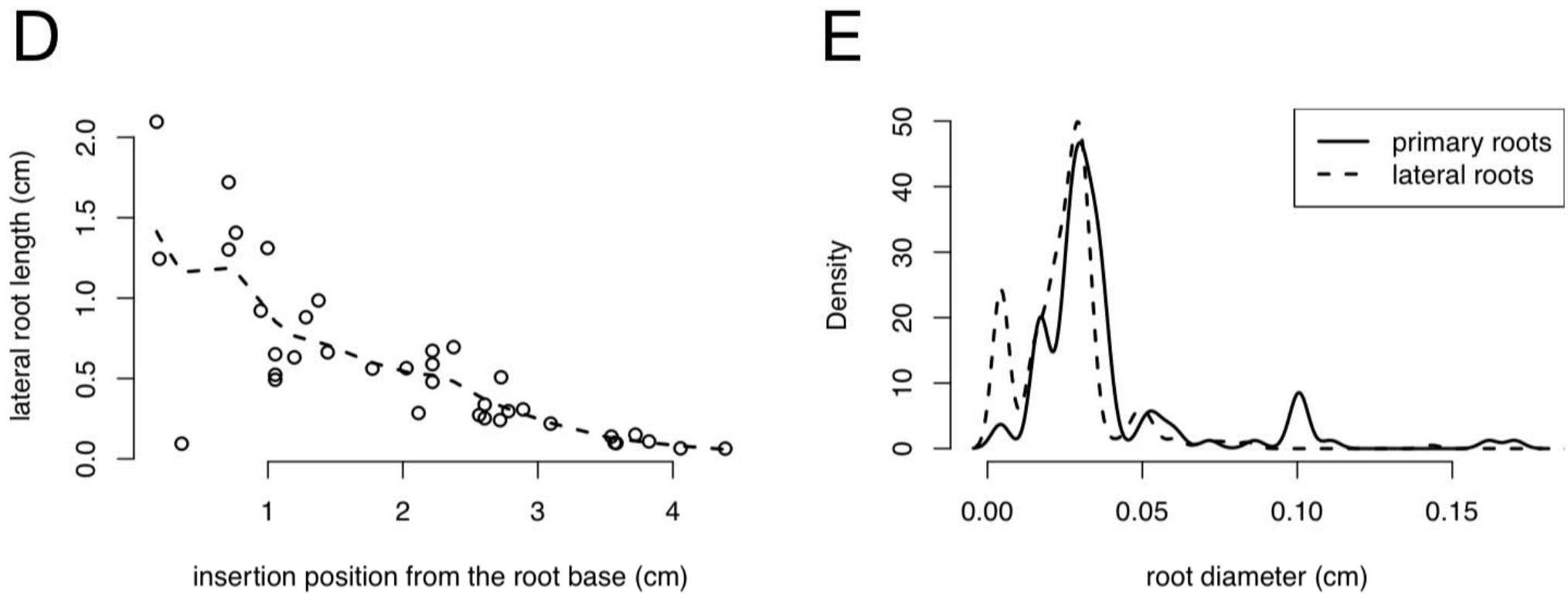
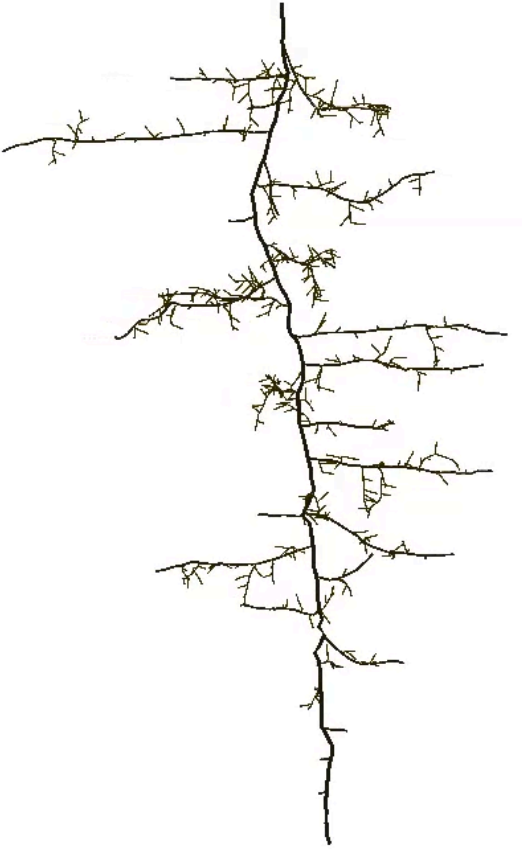


Image Analysis
RooTrak
RootNav
RootTrace
SmartRoot
RootSystem Analyser
RhizoScan
Online Repository
Plant Ontologies
XML
RSML
MTG
RSWMS
ArchiSimple
Root Box
Open Alea
ImageJ
Excell
C#
R
Modeling
Data Analysis

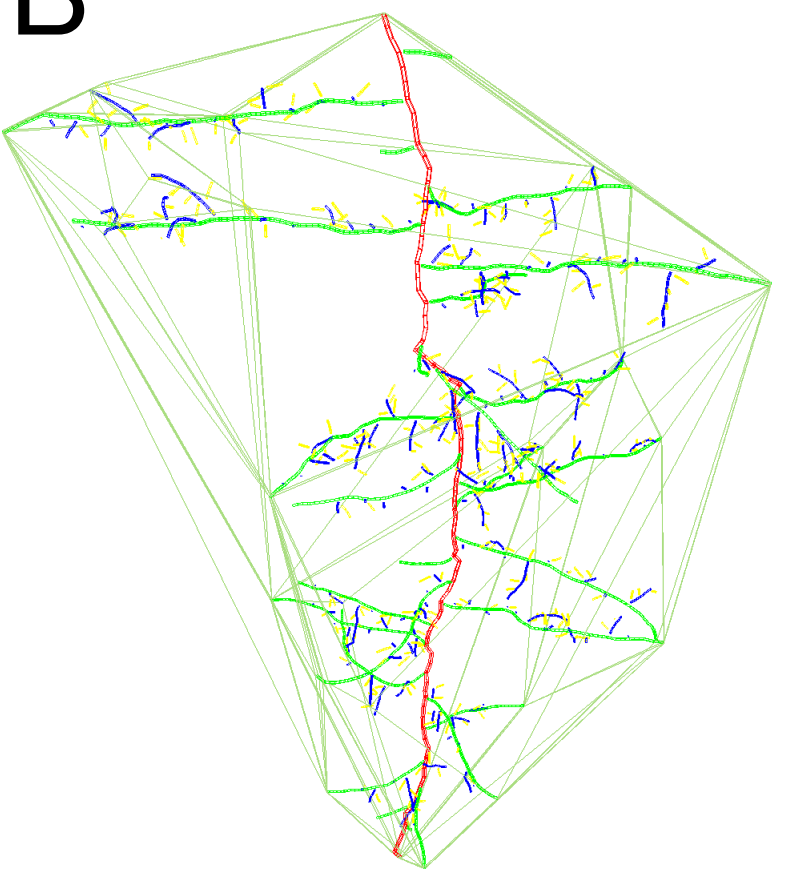
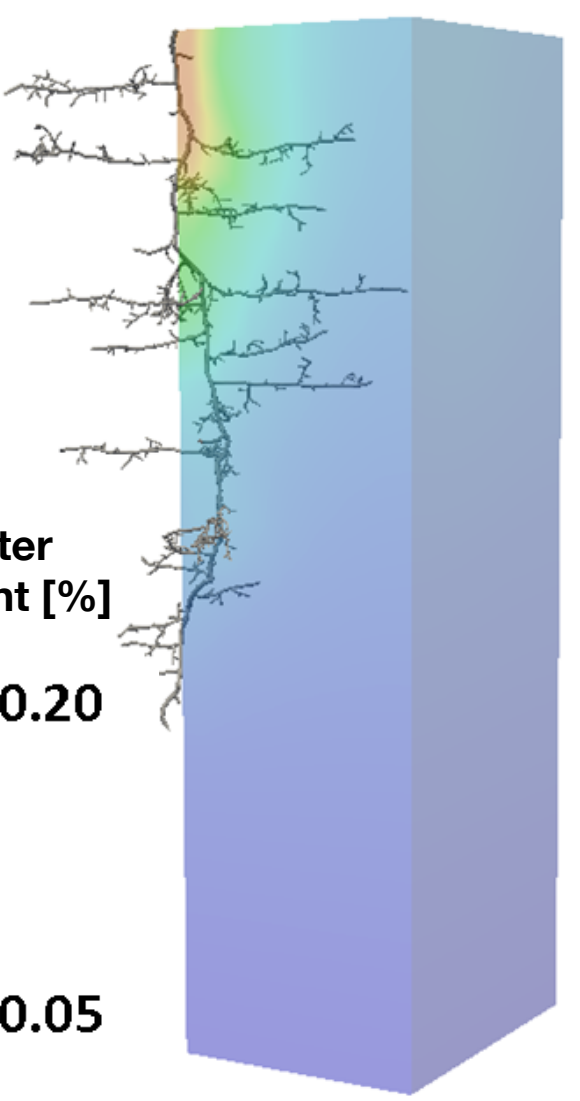

A
B
C


D
lateral root length (cm)
0.0
0.5
1.0
1.5
2.0
1
2
3
4
insertion position from the root base (cm)
E
Density
0
10
20
30
40
50
primary roots
lateral roots
0.00
0.05
0.10
0.15
root diameter (cm)

A



B



C