Project no. 043251

EDEN

ECOLOGICAL DIVERSITY AND EVOLUTIONARY NETWORKS

Instrument: Specific Targeted Project (STREP)

Thematic Priority: Integrating and strengthening the European Research Area
NEST Pathfinder initiative Tackling Complexity in Science

D2.4 (D4): Software Toolbox V2.0

Due date of deliverable: Month 30
Actual submission date: Month 30

Start date of project: 1 January 2007   Duration: 36 months

Organisation name of lead contractor for this deliverable: TKK/LCE

Project co-funded by the European Commission within the Sixth Framework Programme (2002-2006)		
Dissemination Level		
PU	Public	X
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

NOTE: This is a short description of deliverable D2.4 . The deliverable itself is in
`http://ifisc.uib-csic.es/EDEN/fichaOutput.php?idOutput=15`

EDEN Network Analysis Software Toolbox V2
WP2 / Deliverable D2.4
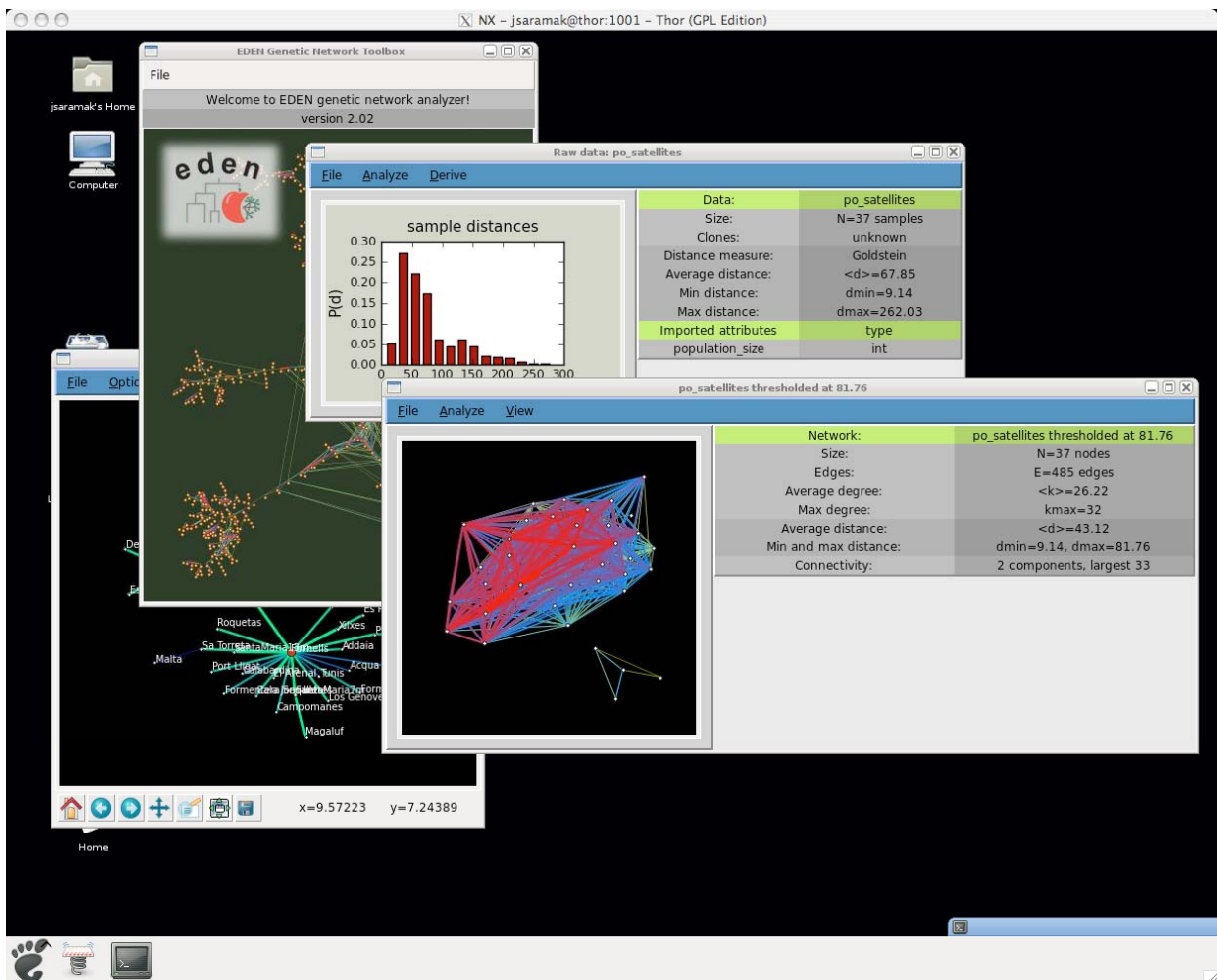
One of the objectives in the EDEN project is to implement the network techniques developed for the analysis of ecological and evolutionary networks into software applications which could be used beyond the frame of the project. The 'Software Toolbox', which has been now more descriptively named **EDEN Network Analysis Software Toolbox** is this intended contribution of EDEN to the scientific community. According to the EDEN workplan, the delivery of the Toolbox would be given in two stages. First, as a software prototype for internal use (deliverable D2.1) and finally, close to the end of the project as version for public use (deliverable D2.4).

The first version of the EDEN Network Analyzer was completed in December 2007 and posted in the restricted access part of the EDEN website 'output' section, more specifically at URL:

http://ifisc.uib-csic.es/EDEN/fichaOutput.php?idOutput=7

The second version (V2.0) of the software package was completed in summer 2009 and posted on the EDEN website 'output' section, at URL:

Since then, improvement and updating of the software has continued, and the current version number is V2.1. The software will continue to be updated in the future as well. An overview of the features of the software toolbox are given below.

EDENetwork network analysis software toolbox – quick overview

The EDENetwork network analysis software toolbox (version 2.1, public version) contains functionalities needed for construction and analysis of genetic distance networks from raw genetic data. The software has been designed considering biologists and genetic researchers as its target users.

The operation of the software follows a logical progression, from importing raw genetic data to construction of genetic distance matrices, deriving thresholded networks of such matrices and analyzing the properties of these networks.

1) **Data importing:** the software reads raw microsatellite repetition data, such that each sampled specimen is simultaneously associated to a sampling location. Optionally, additional data such as geocoordinates or user-defined variables can also be read in.

2) **Genetic distance matrices:** based on user choice, the imported genetic data is transformed either to population-level distances (based on the Goldshtein measure) or distances between individuals (several distance measures: Linear Manhattan, non-shared alleles, etc). At this stage, the user gets a report window on the basic statistical properties of the distance data, including a histogram of distances.

3) **Transformation of matrices to networks:** for quick visual analysis and data exploration, the software allows the user to visualize the data in the form of a Minimal Spanning Tree. For thresholded network analysis, the software first performs percolation analysis where matrix elements are successively removed until the resulting network is fragmented; the user can specify the limit of thresholding based on visual feedback of percolation quantities (largest component size, susceptibility).

4) **Network analysis and visualization:** the resulting networks can be analyzed using standard network analysis techniques (such as calculating various distributions) and exported to files. Furthermore, networks can be visualized (based on a spring-charge-algorithm) such that visualization properties are chosen by the user – nodes and edges can be colored and resized according to sampling locations, user-defined auxiliary data, and network properties such as betweenness centrality. Visualizations can be saved as PNG and EPS files, and are of publication-ready quality.