PharmTree 2.1

M. Karthikeyan*

Information Division, National Chemical Laboratory, Pune - 411 008, India

Received August 16, 2003

PharmTree is a Web-based software package that can fingerprint and classify a large number of chemical compounds according to their biological response data for identifying leads. This program is available from Bio Relation Corporation, 5668 Drysdale Dr., San Jose, CA 95124 (phone: 408-621-6712, e-mail: contact@biorelation.com, http://www.biorelation.com/).

It is strongly recommended that one read the accompanying tutorial to get the overall view and capabilities of the program before trying with one's own data sets. An html version of the table of contents and the user guide are available for a stepwise demonstration on given sample data files. Familiarizing oneself with the processes involved in creating projects with secured access, understanding file formats such as .csv (comma separated values) and SDF (Standard Data Format), and linking the structure data with activity data through common field names, etc. are also recommended. Initial installation of the program requires at least 300 megabytes in the hard drive, as it is integrated with its own Java based Web server to access the program across the network. Users can share the project across the network with secured access. A free 30-day evaluation copy can be downloaded from the company site, and it requires a registration key for installation.[1,2]

Molecules in SDF and MOS format can be used for creating PharmTree where the leadlike molecules can be discovered. With secured access one can create a workspace for a new project. Once the project area is created a user can upload a set of chemical structures in MOS or MDL's SDF file format along with other necessary information such as CAS RN, NSC_ID, or any data identified by data field names. The uploaded molecule is displayed in both 3D and 2D panels for modification and review by the user. In the first step it accepts both chemical structure data in SDF (standard data format) and biological data in CSV (comma separated value) and merges them into one common central-

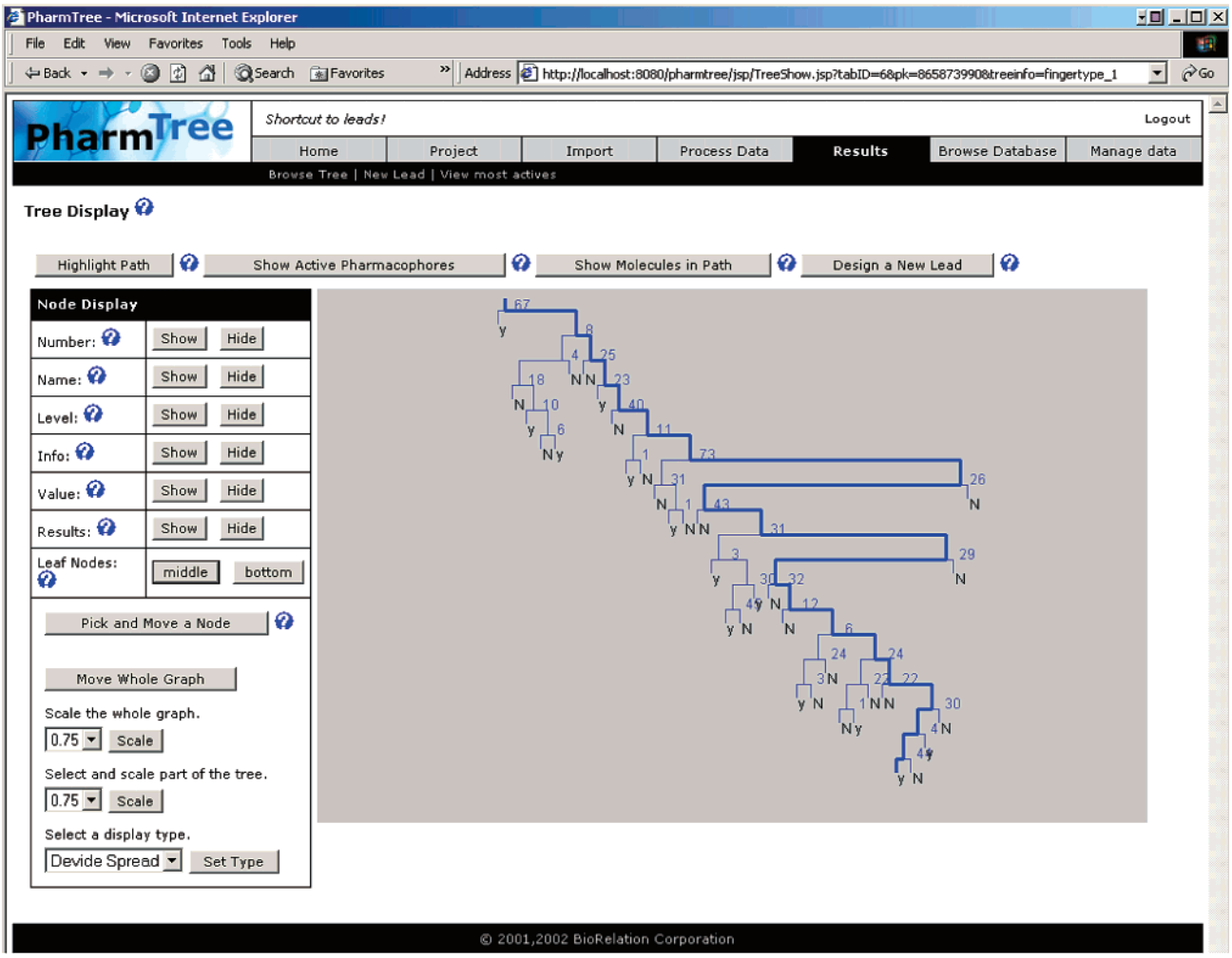

Figure 1. PharmTree results Navigation.

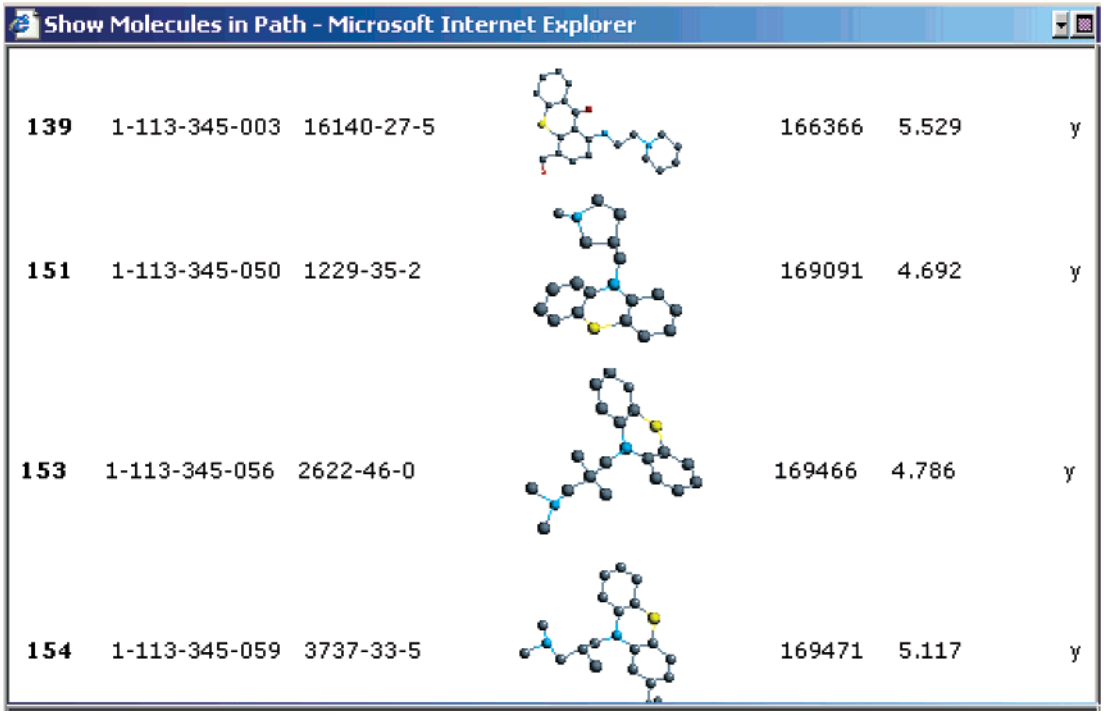

Figure 2. Display Molecules in the path.

ized database. "CSF - Comma delimited Structure File" can be "Uploaded" and mapped against corresponding fields in an activity data file in the database. The use of 2D and 3D fingerprinting in classifying the leads based on its structure connectivity recursively learns from experimental biological activity data such as NLOGG150 to produce a graphical tree based on a chemical classification.[1,2]

The generated pharmtree can be navigated through all tree branches. It will assist in finding the location of active compounds and understanding the active pharmacophores. It also allows users to define rules interactively based on biological activity. One can define its own activity criteria to define or select one's own model (Figure 1).

This new version of PharmTree 2.1 is integrated with PharmPhorePhlex3D technology to classify 3D fingerprints. All pharmacophores are derived from the combinations of Hydrogen Bond Donors (HBD) and Hydrogen Bond Acceptors (HBA) and Hydrophobic and Aromatic centers plus different distance ranges.

•Other additional features of PharmTree 2.1 include 3D viewing of Ball & Stick models.

•Browse the PharmaTree for active molecules in the paths and to identify the contributing pharmacophores. However the time required to display the results depends on the size of the database (Figure 2).

•Design New Lead.

•Export as Excel Sheets in CSV for future analysis and review the results.

All these processes are sequentially arranged for easy operation and are just a mouse click away. During the process of calculation to generate a tree one can view the progress of the results interactively without interrupting the program. Other managerial tools will help to back up the data in a file or delete the whole database to work on a totally new project. PharmTree 2.1 will allow users to classify combinatorial libraries or other large sets of chemical compounds based on the PharmPhoreFlex3D fingerprints. PharmTree is a pharmacophore classifier for biological properties of chemical compounds. It generates large dimensions of pharmacophore spaces based on small drug molecules' 3D conformations. The pharmacophore spaces can be used for a variety of drug relational studies. PharmTree uses the tree based classification technology to classify chemical compounds into different classes. It generates patterns for active compounds. It also relates a variety of biological properties to each other.

PharmTree requires a minimum of a Pentium II Windows-compatible PC running either Windows 98, 2000, or XP. A Pentium 4 with at least 128 MB RAM is highly recommended. This allows the software to run quickly and prevents seizing of the system when carrying out iterated calculations. The software loads from a CD-ROM and takes up 300 MB of hard drive space. A list of sample structural files with a user guide makes full use of the software. PharmTree was tested on HP Pavilion HP Notebook (ZE5375US) with an Intel Pentium 4 Processor 2.4 GHz with 512 MB RAM and a CD-ROM. It used an operating system running Windows XP.

REFERENCES AND NOTES

(1) Binary Formal Inference-Based Recursive Modeling Using Multiple Atoms and Physicochemical Property Class Pair and Torsion Descriptors as Decision Criteria. Cho, S. J. et al. *J. Chem. Inf. Comput. Sci.* **2000**, *40*, 668−680

(2) Structure−Activity Data Set Using Recursive Partitioning. *Quant. Struct. Act. Relat.* **1997**, *16*, 296−302.

* Corresponding author e-mail: karthi@ems.ncl.res.in.

CI034179J