



PedNavigator: a pedigree drawing servlet for large and inbred populations

Gianmaria Mancosu¹, Giuseppe Ledda¹ and Paola M. Melis²

¹Shardna Life Science, Cagliari, Italy and ²Istituto di Genetica delle Popolazioni, CNR, Alghero, Italy

Received on September 9, 2002; revised on October 22, 2002; accepted on November 1, 2002

ABSTRACT

Summary: PedNavigator is a pedigree drawing application for large and complex pedigrees. It has been developed especially for genetic and epidemiological studies of isolated populations characterized by high inbreeding and multiple matrimonyes. PedNavigator is written in Java and is intended as a server-side web application, allowing researchers to 'walk' through family ties by point-and-clicking on person's symbols. The application is able to enrich the pedigree drawings with genotypic and phenotypic information taken from the underlying relational database.

Availability: Available upon request to author.

Contact: mancosu@shardna.it

Manipulating genealogical information is a central task in many genetic and epidemiological studies. Genealogies can be composed of thousands of individuals and typically include loops, multiple mates and many related families. The problem of how to graphically represent all this data is often solved by ignoring some information and simplifying family relationships. While software like Cyrillic, Progeny, GAP, CoPE and GIS Infobiogen, 1999 (Brun-Samarq et al., 1999) are suited to show quite complex structures and large family pedigrees with phenotype and genotype information, they are not suited to contemporaneously show all the existing relationships among the pedigree members.

In our studies of a remote area of Sardinia (Angius et al., 2001), we found that 80–85% of the inhabitants of each village belong to a single huge pedigree with families connected to one another through hundreds of loops. To our knowledge, none of the programs mentioned above is capable of drawing the entire pedigree, compelling researchers to work with subsets of the overall genealogy. Unfortunately, splitting pedigrees into smaller subunits causes a serious loss of information.

In order to give researchers a suitable support to visually explore such a complex pedigree we implemented a graphical application, PedNavigator, as a front-end for genealogical databases. It consists of a servlet, entirely writ-

ten in Java, which can be installed on a Web Application Server and can be consulted through any www graphic browser. The pedigree is graphically represented following Bennett et al. (1995).

PedNavigator accesses a relational database using SQL queries via a standard JDBC Java package, and is able to operate with any database provided with a JDBC driver (e.g. Oracle, MySQL, Microsoft Access, Paradox).

Within the database scheme, each subject has a unique identification code ID, by means of which authorized researchers—following privacy rules—are able to obtain information about subjects under study, for example age, phenotypic data, etc. The main table for storing kinship data between persons is the link table, composed of records with three fields: the individual's, father's and mother's ID.

To accomplish the drawing task, PedNavigator uses a sliding window concept: instead of drawing the whole pedigree, it draws a sub-pedigree centered on a selected individual showing relatives according to certain rules for a given range. The user-friendly interface shown in Figure 1 allows one to select an individual by simply point-and-clicking on its symbolic representation. PedNavigator then redraws the genealogy centering it on the selected person (which appears in red) expanding and showing relations between closest relatives. Proceeding thus, researchers can navigate in real-time through the whole pedigree.

To facilitate the navigation a '+' is drawn near the person's symbol to indicate that the pedigree can continue through that person towards subsequent generations. Within the same generation, the lines linking various persons of each family are drawn in different colors to clearly delineate relations in case of line crossings.

From an architectural standpoint, PedNavigator consists of two software modules: the first module is responsible for consulting the database, taking a person's ID as input and returning a list of related individuals following a given extraction algorithm; the second module takes this list and builds the pedigree layout, arranging persons' symbols vertically according to the generation level and spacing

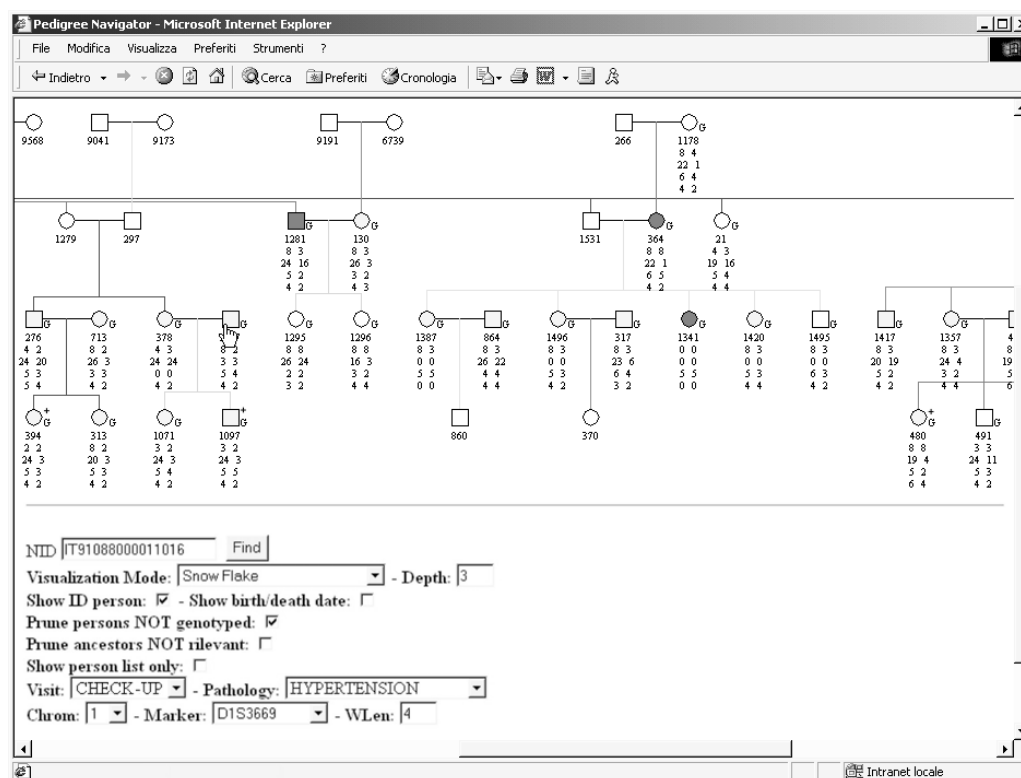


Fig. 1. PedNavigator main page example: dark (blue) symbols show affected persons, gray (yellow) symbols non-affected and white symbols unknown state. Symbols are labeled with selected marker alleles. Plus signs near symbols indicate persons with expandable family (expansion not shown because of setting of expansion limit to 3).

them horizontally to improve clarity.

The database extraction algorithm is configurable to follow only the relevant relationships to limit the number of persons included in any pedigree's representation to maximize graphical clarity. In PedNavigator four extraction methods have so far been implemented:

- (1) Snowflake (SF);
- (2) Direct Ancestors (DA);
- (3) Full Descendent Tree (FDT);
- (4) Direct Paternal Descendants (DPD).

The SF mode includes all persons directly related with the person under study, by following offspring, mating links and parents until a given depth is reached. The DA method acts according to the classic pedigree representation, in which ancestors are shown for a given number of generations. The FDT mode supplies the descendants of an individual, including all mates. The DPD acts like the former but follows descendants according to paternal lineage—useful in studies on the spread of surnames.

With PedNavigator it is possible to enrich each pedigree with further individual information, i.e. genetic data, using

the ID code as a relational key to access database records. For example, researchers can decide to show genetic markers for a family under investigation and also which individual is affected by a given pathology. The output from PedNavigator can be similar to what is shown in Figure 1.

ACKNOWLEDGEMENTS

We thank Shardna and Institute of Population Genetics colleagues for their suggestions and the Diocese of Lanusei for access to their parochial historical archives.

REFERENCES

- Angius,A., Melis,P.M., Morelli,L., Petretto,E., Casu,G., Maestrale,G.B., Fraumene,C., Bebbere,D., Forabosco,P. and Pirastu, (2001) Archival, demographic and genetic studies define a Sardinian sub-isolate as a suitable model for mapping complex traits. *J. Hum. Genet.*, **109**, 198–209.
- Bennett,R.L. *et al.* (1995) Recommendations for standardized human pedigree nomenclature. *Am. J. Hum. Genet.*, **56**, 745–752.
- Brun-Samarq,L., Gallina,S., Philippi,A., Demenais,F., Vaisseix,G. and Barillot,E. (1999) CoPE: a collaborative pedigree drawing environment. *Bioinformatics*, **15**, 345–346.