BIOINFORMATICS APPLICATIONS NOTE Vol. 23 no. 14 2007, pages 1854–1856
doi:10.1093/bioinformatics/btm242

Genetics and population analysis

Madeline 2.0 PDE: a new program for local and web-based pedigree drawing

Edward H. Trager[1,*,†], Ritu Khanna[1,†], Adrian Marrs[1,2], Lawrence Siden[3], Kari E.H. Branham[1], Anand Swaroop[1] and Julia E. Richards[1]

[1]Department of Ophthalmology and Visual Sciences, Kellogg Eye Center (KEC), University of Michigan, Ann Arbor, Michigan 48105, [2]Terrain Experts, Inc. (TERREX), Tucson, Arizona 85711 and [3]Independent Consultant, Ann Arbor, Michigan 48197, USA

Received on March 1, 2007; revised on April 10, 2007; accepted on April 29, 2007

Advance Access publication May 8, 2007

Associate Editor: Martin Bishop

ABSTRACT

Summary: The Madeline 2.0 Pedigree Drawing Engine (PDE) is a pedigree drawing program for use in linkage and family-based association studies. The program is designed to handle large and complex pedigrees with an emphasis on readability and aesthetics. For complex pedigrees, we use a hybrid algorithm in which consanguinous loops are drawn as cyclic graphs whenever possible, but we resort to acyclic graphs when matings can no longer be connected without line crossings. A similar hybrid approach is used to avoid line crossings for matings between distant descendants of different founding groups. Written in object-oriented C++ and released under the GNU General Public License (GPL), Madeline 2.0 PDE reads input files specified on the command line and generates pedigree drawings without user interaction. Pedigree output in scalable vector graphics (SVG) format can be viewed in browsers with native SVG rendering support or in vector graphics editors. We provide an easy-to-use public web service, which is experimental and still under development.

Availability: http://kellogg.umich.edu/madeline

Contact: ed.trager@gmail.com

Clinical genetic studies often involve extended families containing tens or hundreds of individuals. Consanguinous loops, multiple mated individuals and several founding groups can occur in such pedigrees. Manual data entry of these pedigrees is both time consuming and error prone. As demonstrated by Tores and Barillot (2001), finding a solution for drawing complex pedigrees is a challenging task. Most available programs either display a subset of the relationships among pedigree members or ignore the aesthetic criteria that govern the optimal visualization of ordered trees (Mäkinen *et al*., 2005; Mancosu *et al*., 2003).

Madeline v. 0.935 (Trager, 2004), a program for managing and drawing pedigrees, continues to be widely used by researchers around the world (Ching *et al*., 2006; Dudbridge *et al*., 2004; Shimizu *et al*., 2004). However its pedigree drawing routines have many limitations because the program was originally designed to only handle single descent trees of simple nuclear families. Four other freely available pedigree drawing programs that implement at least some of the recommendations of the Pedigree Standardization Task Force (PSTF, Bennett *et al*., 1995)—Cranefoot (Mäkinen *et al*., 2005), Haplopainter (Thiele and Nürnberg, 2005), Kinship (Zhao, 2006) and PedigreeQuery (Kirichenko, 2004)—are also limited when drawing complex pedigrees. For example, Haplopainter (Thiele and Nürnberg, 2005) uses a non-traditional approach to draw multiple consanguinous loop pedigrees. A detailed comparison is available on the Madeline web site. To overcome these limitations, we have developed the Madeline 2.0 Pedigree Drawing Engine (PDE) as a program, which does not compromise on the aesthetics or readability of a pedigree and is capable of preparing and visualizing complex pedigrees. We additionally created a web-based experimental service, which researchers can use to create, display, save and print pedigrees. Detailed documentation is available on the web site.

Madeline 2.0 PDE implements a subset of the recommendations of the PSTF (Bennett *et al*., 1995), where square icons represent males and circular icons females. The orthogonal-segment graph style espoused by PSTF is simple and visually appealing; however, algorithmic issues arise with complex pedigrees. Tores and Barillot (2001) define the concept of a perfectly drawable pedigree (PDP) and demonstrate that the existence of a perfect solution can be tested in linear time and be drawn if the pedigree is PDP. When a pedigree is not PDP, the solution remains non-trivial. The approach used by Mäkinen *et al*. (2005) involves transforming the pedigree into an acyclic graph.

We use a hybrid algorithm in which consanguinous loops (CLs) are drawn as cyclic graphs, whenever possible (Fig. 1). We resort to acyclic graphs when matings can no longer be connected without line crossings. We apply a similar approach to avoid line crossings in matings between distant descendants of different founding groups (DFGs). In both cases, we reorder siblings within nuclear families so that mated individuals are as close to their respective CL or DFG spouses as possible.

*To whom correspondence should be addressed.

†The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

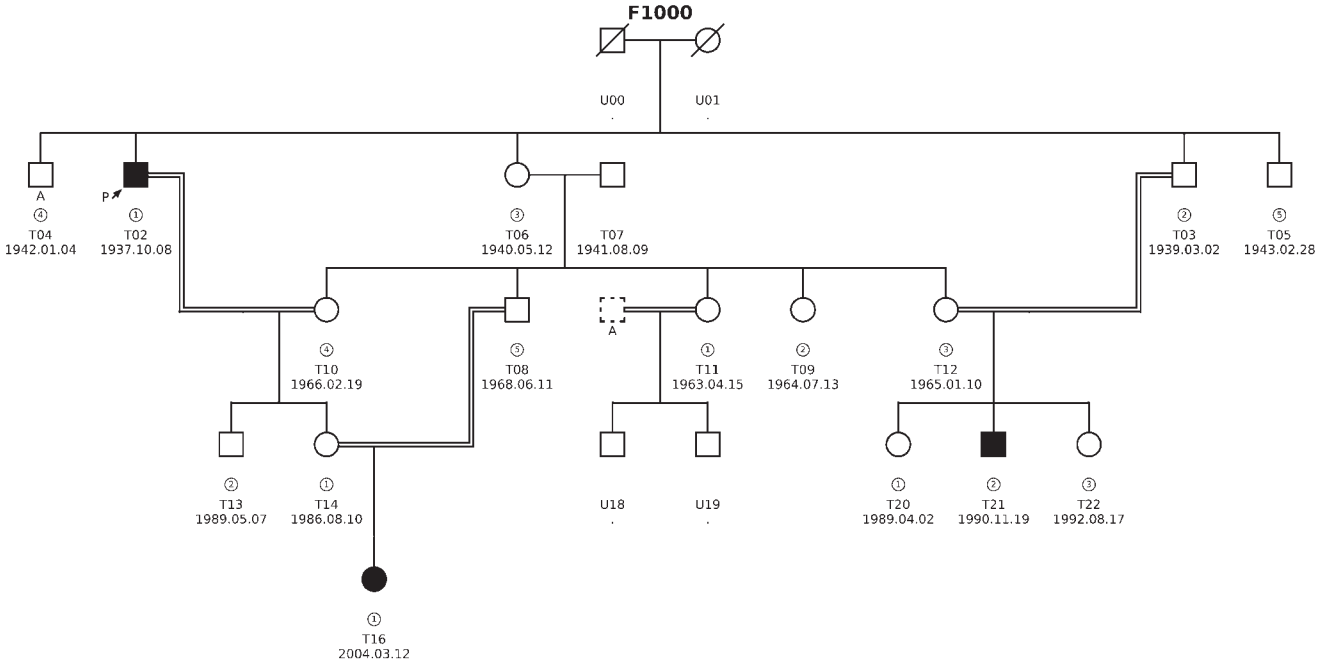

Fig. 1. Sample pedigree with four CL matings. Direct connecting lines are drawn for three of the four CL matings. However, the fourth CL mating would require a line crossing; hence, individual T04 is shown at a second location (dashed box) and marked directly below both icons with a unique label 'A' to indicate the repeat. Circled numbers indicate birth order within sibships. Shaded icons represent affected individuals. Individual identifiers and dates of birth are also displayed.

This approach ensures that within any given nuclear family, up to two reordered CL or DFG-mated siblings—the leftmost and the rightmost, respectively—can have direct non-crossing lines connecting to their spouses. Spouses in remaining CL or DFG matings are drawn using dashed icons at additional locations on the graph, as required to avoid crossing lines. Multiple mates and twin groups can constrain sibling reordering, preventing the drawing of non-crossing lines connecting spouses. Small circled numbers can be displayed to show birth order in reordered sibships.

Pedigree construction minimally requires family, individual, mother, father and gender identifiers. Additional phenotype and genotype columns can be included in the data file. Display columns can be specified on the command line or in a separate file. Column names starting with 'Affected' are automatically used to shade icons.

Madeline 2.0 PDE supports flat, XHTML, OASIS, Office Open XML, Madeline XML and, optionally, MySQL file formats read from disk or using HTTP/S network protocols. A single data file may contain multiple pedigrees. UTF-8 Unicode data and international digits are supported. Phenotype data may contain imprecise values in the form of approximations or ranges. Full support for search queries against ranged and approximate data is planned for a future interactive version.

When constructing pedigrees, Madeline 2.0 checks for common errors, such as gender mismatches between mated pairs and non-founders missing a single parent. To streamline data processing for large families, the program has been designed to issue warning messages and take corrective action whenever possible.

We have integrated Madeline 2.0 into an easy-to-use web-based service, which can greatly simplify the process of creating a pedigree data file. The public service uses AJAX and embedded SVG: a browser capable of rendering SVG natively (Firefox 1.5+, Opera 9.0+) is required. The command-line '–embedded' option facilitates integrating Madeline 2.0 SVG output into web-based applications.

Madeline has been applied in age-related macular degeneration (AMD) and primary open angle glaucoma (POAG) linkage studies at KEC (Abecasis *et al*., 2004; Woodroffe *et al*., 2006). Active development and use of the program continues. Madeline v. 0.935, the predecessor of Madeline 2.0 PDE, converts pedigree and marker data into various formats required by linkage analysis software and also provides functionality for querying pedigree data sets interactively. These features will also be incorporated into future releases of Madeline 2.0 PDE.

ACKNOWLEDGEMENTS

We would like to thank Alyce M. Dohring, Mohammad Othman and David M. Reed for providing us with useful feedback and sample data. This work was supported by a generous grant from the Elmer and Sylvia Sramek Foundation, and in part by from the National Institutes of Health (EY001167, EY016862, EY007003), The Foundation Fighting Blindness, Research to Prevent Blindness, and the Van Arnum Glaucoma Research Fund.

Conflict of Interest: none declared.

REFERENCES

Abecasis,G.R. *et al.* (2004) Age-related macular degeneration: a high-resolution genome scan for susceptibility loci in a population enriched for late-stage disease. *Am. J. Hum. Genet.*, **74**, 482–494.

Bennett,R.L. *et al.* (1995) Recommendations for standardized human pedigree nomenclature. *Am. J. Hum. Genet.*, **56**, 745–752.

Ching,K.A. *et al.* (2006) Data and animal management software for large-scale phenotype screening. *Mamm. Genome.*, **17**, 288–297.

Dudbridge,F. *et al.* (2004) Pelican: pedigree editor for linkage computer analysis. *Bioinformatics*, **20**, 2327–2328.

Kirichenko,A.V. (2004) An algorithm of step-by-step pedigree drawing. *Russ. J. Genet.*, **40**, 1425–1428.

Mancosu,G. *et al.* (2003) PedNavigator: a pedigree drawing servlet for large and inbred populations. *Bioinformatics*, **20**, 669–670.

Mäkinen,V.-P. *et al.* (2005) High-throughput pedigree drawing. *Eur. J. Hum. Genet.*, **13**, 987–989.

Shimizu,S. *et al.* (2004) A locus for posterior polymorphous corneal dystrophy (PPCD3) maps to chromosome 10. *Am. J. Med. Genet.*, **130A**, 372–377.

Thiele,H. and Nürnberg,P. (2005) Haplopainter: a tool for drawing pedigrees with complex haplotypes. *Bioinformatics*, **21**, 1730–1732.

Tores,F. and Barillot,E. (2001) The art of pedigree drawing: algorithmic aspects. *Bioinformatics*, **17**, 174–179.

Trager,E.H. (2004) Madeline v. 0.935 documentation. http://eyegene.ophthy.med.umich.edu/madeline-0.935/.

Woodroffe,A. *et al.* (2006) Ordered subset analysis supports a glaucoma locus at GLC1I on chromosome 15 in families with earlier adult age at diagnosis. *Exp. Eye Res.*, **82**, 1068–1074.

Zhao,J.H. (2006) Pedigree-drawing with R and graphviz. *Bioinformatics*, **22**, 1013–1014.