

Pedigree online: A New Web Server for Two-Way Interactive Pedigree Drawing

Ye Chen^{1*}, Xiaohui Cang^{2,3*}

¹School of Information and Control Engineering, China University of Mining and Technology, Xuzhou, Jiangsu 221008, China

²Division of Medical Genetics and Genomics, The Children's Hospital, Zhejiang University School of Medicine

³Institute of Genetics, Zhejiang University and Department of Genetics, Zhejiang University School of Medicine, Hangzhou, Zhejiang 310058, China

*Corresponding author

Xiaohui Cang, Division of Medical Genetics and Genomics, The Children's Hospital, Zhejiang University School of Medicine and Institute of Genetics, Zhejiang University and Department of Genetics, Zhejiang University School of Medicine, Hangzhou, Zhejiang 310058, China.

Submitted: 07 Nov 2021; **Accepted:** 13 Dec 2021; **Published:** 23 Dec 2021

Citation: Ye Chen, Xiaohui Cang. (2021). PedigreeOnline: a new web server for two-way interactive pedigree drawing. *J Gene Engg Bio Res*, 3(2), 64-68.

Abstract

Pedigree drawing is of special importance for genetic counseling and research on documenting families with inherited disease. Current pedigree-drawing tools fall into two types: those build pedigree diagrams from input family data, and those provide a graphical interface to draw pedigree directly. Here we merge the two types of programs into one, and release a brand new pedigree-drawing web server, PedigreeOnline. Through introducing a novel pedigree-like-layout (PLL) family data format and taking advantage of HTML5 canvas element in development, this program is able to build two-way interactive connections between family data and graphical interface. Users can work on and switch easily between the graphical interface and the family data window. Currently this web server is compatible with Linkage data format and comma-separated value (CSV) data format. These features make PedigreeOnline a pedigree-drawing tool with outstanding user-friendliness and flexibility. PedigreeOnline is freely accessible for everyone at: <http://www.ig.zju.edu.cn/PedigreeOnline/>

Keywords: Pedigree Drawing, Web Server, Two-Way Communication, Compatibility, Linkage Data Format

Introduction

Collection, update and visualization of family history (pedigree data) are of crucial importance for genetic counseling and genetic research on families with inherited diseases. A variety of pedigree-drawing tools including stand-alone and web-based programs have been developed. These programs fall into two classes: those that generate pedigree diagrams from the input family data, and those that provide graphical interface for users to draw pedigree directly.

For the first type of programs such as Madeline 2.0 and CraneFoot, family data is stored and manipulated in a multi-column table, in which each line is used to describe information for each family member, including healthy ones. However, only limited information for healthy members is needed to build a pedigree, and healthy members usually constitute the vast majority in most families under study [1, 2]. Moreover, even for the information such as proband status, the table needs one column to label it although there is only one proband member in the whole family. Therefore, such a multi-column table dramatically increases the size and decrease

the readability of the data file, and manual creation of these files therefore becomes time consuming and error prone.

For the second type of programs, drawing on the graphical interface improves the user-friendliness, but the users have no option to save their state files in some programs, or in some other programs they have to save their graphical state files within the registered account, which makes the future update inconvenient. Some recent programs such as pedigreejs and panogram (<http://panogram.github.io/panogram/>) made improvements and allow users to draw pedigree on the graphical interface and export the family data file for future update [3]. But these data files are designed only for the programs and not readable for users, therefore users have to import the data file back into the program and work on graphical interface for any change or update. In many cases, another data set understandable for users still needs to be maintained separately.

A recent survey showed that nearly half of cancer genetic counselors (CGCs) were considering adopting or changing their pedigree drawing tool, and meanwhile they were concerned about budget

and compatibility of the new tool to their old data [4, 5]. Therefore, there is a need to provide freely-accessible pedigree-drawing tools with satisfying user-friendliness and compatibility for the geneticists and counselors. The difference in the structure of the family data in which each line represents an individual and the structure of the pedigree diagram in which each line represents one generation forms an obvious obstacle between pedigree datasets and diagrams, leading to different limitations of the two types of programs as mentioned above. Here we overcome these limitations through introducing a novel family data format and merging the two types of programs into one, and release a brand new pedigree-drawing web server, PedigreeOnline.

The Pedigreeonline Web Server

PedigreeOnline is a lightweight pedigree-drawing web server of considerable user-friendliness. It is written in JScript, and works with any web-browser supporting HTML5. PedigreeOnline provides a convenient and free access to all final users, and no registration is needed. PedigreeOnline is featured with a novel family data format, two-way interactive mode between family data and pedigree, and compatibility to other family data formats.

The Pedigree-Like-Layout (PLL) Family Data Format

PedigreeOnline has two windows: family data window and pedigree drawing window. To show the family data concisely within the on-screen window and make data easy for editing, we have introduced a novel pedigree-like-layout (PLL) family data format. This PLL format is absolutely different from any other previously reported family data format in which each line represents one in-

dividual. In the pedigree-like-layout format, as its name suggests, family information is arranged with plain text following the structure of a pedigree, that is, different generations are arranged vertically, and information of the same generation is depicted in one line. In the PLL format (Figure 1A), from top to bottom, each line starting with the signal “#” stands for one generation. A typical generation line is composed of kinship information and phenotype information. The kinship information lists all members of this generation in this way: siblings from the same parent are placed within one parenthesis, with their parents names (father’s name on the left, mother’s name on the right) designated before the “:” sign. If the parents are close relatives, their names are connected with “-”, otherwise with “+”. The kinship information and phenotype information are separated with the “/” sign. In phenotype part, individuals with certain phenotypes are explicitly listed after the corresponding phenotype code (Figure 1B). Different phenotype fields are separated with the “;” sign. After the generation data block, remark information is given after the “*” sign. Such a novel description method drastically decreases the size of the data, and the family data can be easily shown and edited within the on-screen window. Moreover, because family relationship is represented in a more straightforward way, data readability is effectively improved. Because the family data is small and the data arrangement directly reflects the pedigree layout, updating information for an existed family such as adding a new member or changing siblings order can be easily accomplished. In contrast, in the traditional multi-column table, updating family information such as adding a new member needs to be very careful, especially when the new entry is far from her/his parents and siblings in space.

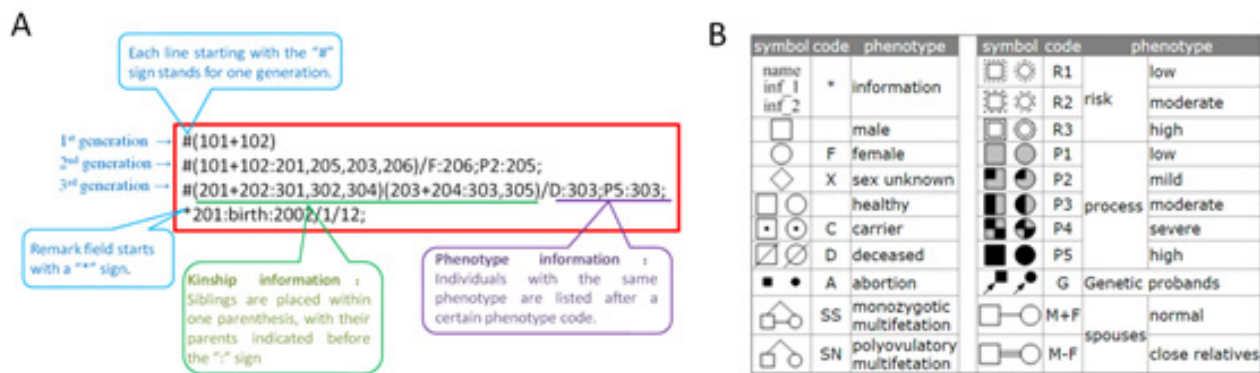


Figure 1: The Pedigree-Like-Layout (PLL) Family Data Format. (A) An example of pedigree data in PLL format is shown within the red block, with brief explanations. (B) The phenotype codes used in the PLL family data and the corresponding symbols applied in the pedigree diagrams.

Two-Way Interactive Connections between Family Data and Pedigree Diagram

Pedigree Online has two major windows: the family-data window and the graphical diagram window. Through introducing the concise and friendly PLL family data format and taking advantage of the HTML5 canvas element in development, we build two-way interactive connections between the family data and the graphi-

cal interface (Figure 2). Users could type or edit the PLL family data within the family-data window, and build a pedigree diagram through a simple click. They could also work on the graphical interface, and any change on the pedigree diagram causes synchronized change in the family data. Users could switch very easily between these two windows at any time when drawing a pedigree.

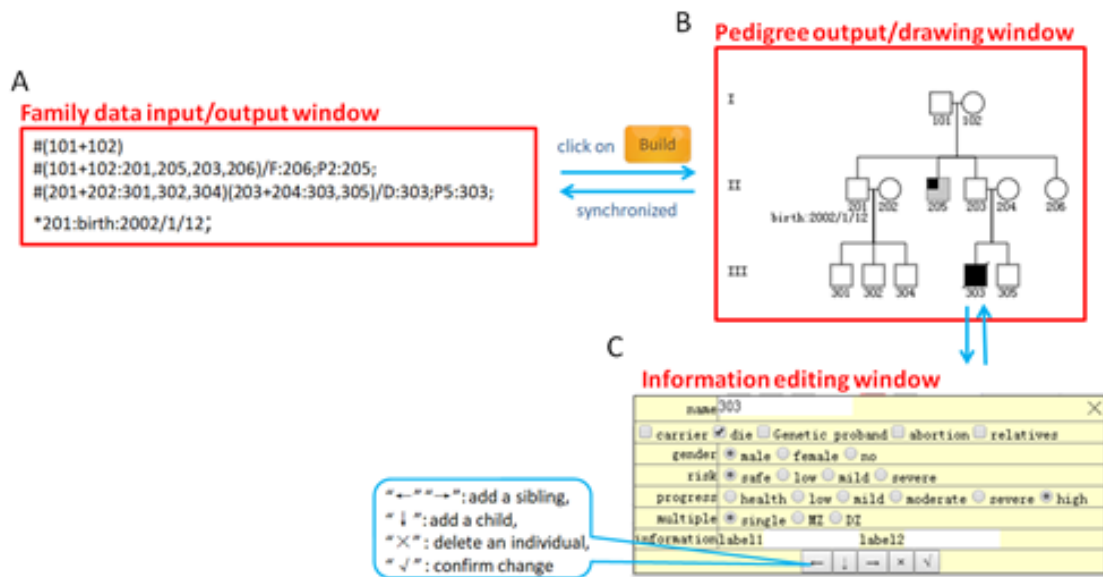


Figure 2: The Two-Way Interactive Mode Between Family Data And Pedigree Diagram. (A) The family data input and output window. (B) The pedigree output and drawing window. (C) The symbol-editing window.

The Algorithm for Symbol Layout in Pedigree Diagram

In Pedigree Online, the symbol layout is calculated and adjusted anytime when a new member is added. To get a beautiful symbol layout, we have used the algorithm shown in Figure 3. In the first step, positions of symbols of the latest generation are calculated (Figure 3A). These symbols are grouped by parents and arranged in the order of birth from left (youngest) to right (oldest), with equal space leaving to every symbol. Above the siblings in the same family, horizontal lines and vertical lines connects these symbols. Calculate the center of horizontal (sibling) lines in same family, based on which symbols positions of their parents can be

determined. In the second step, positions of symbols of the previous generation are calculated. Those individuals without children are added, and the siblings in this generation are also arranged in the order of birth from left (youngest) to right (oldest). Positions of the added symbols are determined by the position of the first parent symbol with children. If there is enough space for the added individuals, no movement is necessary (Figure 3B). Otherwise some parents together with their children need to move the same distance to the right to avoid any overlap (Figure 3C). Then repeat the the second step, until the first generation.

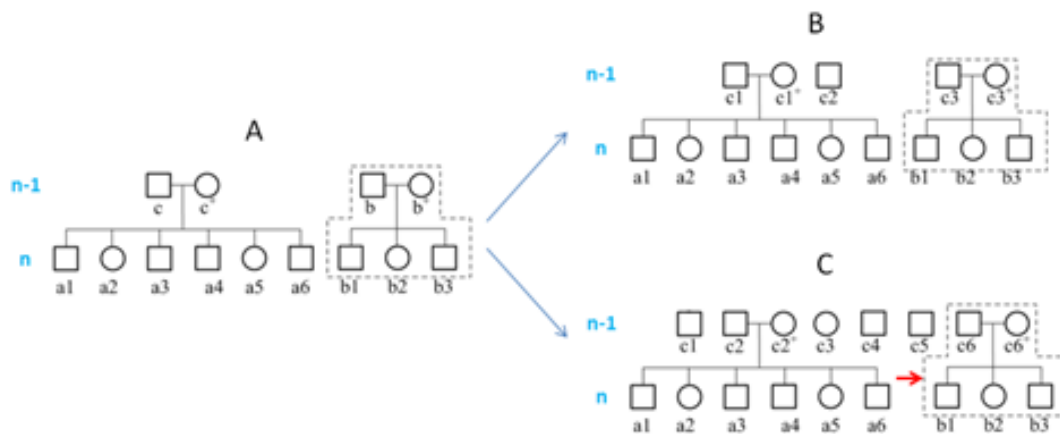


Figure 3: Schematic Illustration on the Algorithm for Symbol Layout. (A) The symbol positions of the latest generation (the nth generation in this case) are first determined, and then symbol positions of their parents (temporarily labeled as: c,c⁺;b,b⁺) are determined accordingly. (B) In the previous generation (the (n-1)th generation), add those individuals without children (c₂ in this case), and number each individual. As there is enough space for the added symbols, no movement is necessary. (C) When there is not enough space for the added individuals (c₁,c₃,c₄,c₅), the sub-family (c₆,c₆⁺,b₁,b₂,b₃) needs to move as a group to the right to avoid any overlap.

Compatibility with Other Family Data Formats

In addition to this new PLL family data format, the common Linkage format and comma-separated value (CSV) format can also be loaded into PedigreeOnline for pedigree drawing, after transforming to the PLL format through the interface provided. Moreover, due to the significance of the Linkage format in genetic analysis, PedigreeOnline also provides a function to transform our PLL data to the Linkage format. Such a two-way transform function dramatically improves the flexibility and compatibility of the software.

Results and Discussion

Input

In PedigreeOnline, users have various ways to input their family data and draw a pedigree.

Type In or Paste Pll Family Data In the On-Screen Window

The PLL format is very concise and easy to follow, so users can easily type in the PLL family data within the on-screen data window (Figure 2A). They can also copy and paste their previously saved PLL data to the on-screen data window for update.

Draw Pedigree Directly on the Graphical Interface

By utilizing the canvas element in HTML5, drawing on graphical interface becomes very easy. Click on any symbol in the pedigree diagram (Figure 2B), a symbol editing window (Figure 2C) will pop up, where users can set up or update information of an individual. Siblings or children can be added through simply clicking on the left/right or downward arrows at the bottom of the window. Deleting an individual from a pedigree can be simply fulfilled by clicking on the “x” icon on the window. When adding the first child for an individual, the program will automatically generate a spouse of the opposite gender.

Load an Existed File in CSV Format

For any existed family data saved in the multi-column table, users

need to modify it according to the format provided in our website, and then save it as CSV file before loading it to the PedigreeOnline. The loaded CSV data can be easily transformed to the PLL format and then used to build pedigree.

Load an Existed File in the Common Linkage Format

Following the suggestions of the editor after the initial evaluation, we have added the function to allow users to load the family data in the common Linkage format. Users can import the Linkage file, and transform it into the PLL data before building the pedigree diagram.

Output

As mentioned above, because of the two-way communications, the family data window and the graphical diagram interface can serve as both input and output windows. In PedigreeOnline, users can easily generate a beautifully arranged pedigree diagram. The size of the graphical window adapts automatically to the size of the pedigree diagram, and if the diagram goes beyond the screen, a scrollbar will appear at the bottom of the screen. The symbols in the diagram follow the recommendations of the PSTF and all the symbols supported by the current version are listed in the Figure 1B. In addition to building a standard pedigree in which each individual is labeled with an Arabic number (Figure 4A), users can also create a pedigree with all individual names shown (Figure 2B), for the convenience of the in-house research [6]. If the family data is transformed from the Linkage format, the genotype information in the Linkage data is saved in the first line of remark field in the PLL data, and users have options to explicitly show the genotype information on the pedigree diagram or not. This operation can be made to all individuals or specifically to one individual. If the genotypes or other remark information of all or some individuals are shown on the pedigree diagram, the space between symbols will be adjusted accordingly (Figure 4B). The pedigree diagram can be saved as a PNG figure of publication-quality.

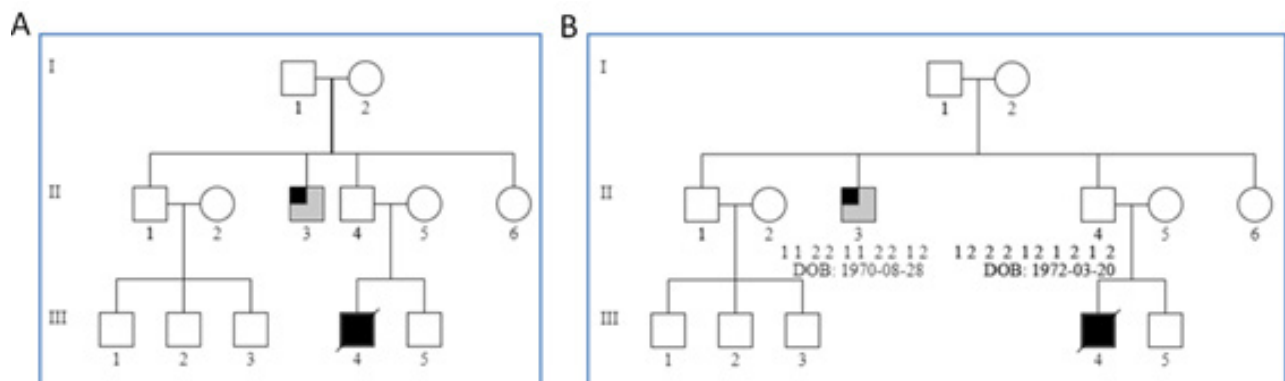


Figure 4: An Output Example of a Standard Pedigree Diagram. (A) An output example of a standard pedigree diagram without remark information. (B) The layout of the same pedigree is adjusted accordingly if the remark information (genotype and date of birth in this case) are explicitly shown on the diagram.

When pedigree is drawn or edited within the graphical interface, the family data updates automatically, therefore when the pedigree drawing is finished, the updated family data need to be saved. As the PLL data is in plain text, it can be saved and edited anywhere

as users like.

Due to the importance of the Linkage data in genetic analysis, PedigreeOnline also allow users to transform the updated PLL family

data to the Linkage format.

Limitations and Future Directions

The current version of PedigreeOnline still has some limitations. The resolution of the individual number or names shown in the diagram needs to be improved. Users can only draw a pedigree from the first generation downward to the later generations, and can not add a new member in a reverse direction. Therefore, only the family history from either paternal part or maternal part of a proband can be drawn. Complicated situations such as consanguineous loops can not be shown in our diagram. We plan to overcome these limitations in the next version of PedigreeOnline. Efforts will also be made to improve the graphical editing ability of the pedigree window, for example, allowing users to select and extract a subgroup from a big pedigree.

Conclusion

Through introducing a novel pedigree-like-layout (PLL) family data format and taking advantage of the HTML5 canvas element in development, the new pedigree-drawing web server, PedigreeOnline accomplishes a two-way interactive mode between family data and graphical interface. In this way, drawing and updating a pedigree diagram become extremely easy and flexible. In PedigreeOnline, the PLL family data format and the common Linkage format can be transformed reciprocally, which will make the software more useful and popular among a broad group of geneticists and counselors.

References

1. Trager, E. H., Khanna, R., Marrs, A., Siden, L., Branham, K. E., Swaroop, A., & Richards, J. E. (2007). Madeline 2.0 PDE: a new program for local and web-based pedigree drawing. *Bioinformatics*, 23(14), 1854-1856.
2. Mäkinen, V. P., Parkkonen, M., Wessman, M., Groop, P. H., Kanninen, T., & Kaski, K. (2005). High-throughput pedigree drawing. *European Journal of Human Genetics*, 13(8), 987-989.
3. Carver, T., Cunningham, A. P., Babb de Villiers, C., Lee, A., Hartley, S., Tischkowitz, M., & Antoniou, A. C. (2018). pedigeejs: a web-based graphical pedigree editor. *Bioinformatics*, 34(6), 1069-1071.
4. Ritchie, J. B., Allen, C. G., Morrison, H., Nichols, M., Lauzon, S. D., Schiffman, J. D., & Welch, B. M. (2020). Utilization of health information technology among cancer genetic counselors. *Molecular Genetics & Genomic Medicine*, 8(8), e1315.
5. Allen, C. G., Ritchie, J. B., Morrison, H., Lauzon, S. D., Nichols, M., Schiffman, J. D., & Welch, B. M. (2021). A thematic analysis of health information technology use among cancer genetic counselors. *Journal of Genetic Counseling*, 30(1), 170-179.
6. Bennett, R. L., Steinhaus, K. A., Uhrich, S. B., O'Sullivan, C. K., Resta, R. G., Lochner-Doyle, D., & Hamanishi, J. (1995). Recommendations for standardized human pedigree nomenclature. *Journal of genetic counseling*, 4(4), 267-279.

Copyright: ©2021 Xiaohui Cang, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.