

A-DNA Utility

© 2014 Charles R. Warthen & T. Wesley Erickson



Autosomal comparison results for Thomas Wesley Erickson

File Edit Settings Help

459 matches for 511 records 5.0 Do It!

Name	Ch	Start Loc	End Loc	cM	SNPs
Pam Bodwell	1	5118461	12148559	12.88	2085
Branda Wohlgemen	1	5118461	10595681	9.29	1596
Hayedeh Sepold Scott	1	19695400	26593874	9.24	2000
Carsten Edmund Malmberg	1	31908360	38625611	8.67	1665
Per-Olof Vilhelm Jonasson c/o Viktori	1	61885351	76790013	14.86	3700
Peter Ferdinand Reid	1	69857385	81219801	8.26	2600
James Lee Lowen	1	70453049	81219801	7.82	2500
Kjell Olof Yngve Yngvesson	1	89157869	101785313	11.1	3700
Scott Gregory Brookover	1	94382066	103926415	8.13	2800
James Spiegelberg	1	94382066	108398491	12.46	3800
Brenda Booth Decristina	1	94841509	106026345	9.39	3100
Amy B Siegle	1	95069962	105550721	8.65	2900
Susan C Griswold	1	96069014	115699858	19.82	5500
Peter Nelson	1	97081292	110902085	14.06	3800
Lonnie Gene Seawright	1	97459480	107100064	8.17	2500



ADNA

Utility for processing autosomal data files

by Charles R. Warthen & T. Wesley Erickson

Copyright © 2014 Charles R. Warthen and T. Wesley Erickson

A-DNA Utility

© 2014 Charles R. Warthen & T. Wesley Erickson

All rights reserved. No parts of this work may be reproduced in any form or by any means - graphic, electronic, or mechanical, including photocopying, recording, taping, or information storage and retrieval systems - without the written permission of the publisher.

Products that are referred to in this document may be either trademarks and/or registered trademarks of the respective owners. The publisher and the author make no claim to these trademarks.

While every precaution has been taken in the preparation of this document, the publisher and the author assume no responsibility for errors or omissions, or for damages resulting from the use of information contained in this document or from the use of programs and source code that may accompany it. In no event shall the publisher and the author be liable for any loss of profit or any other commercial damage caused or alleged to have been caused directly or indirectly by this document.

Printed: September 2014 in Warner Springs, CA.

Publisher

T W Erickson Enterprises

Managing Editor

Jordan A. Ramirez

Technical Editors

Charles R. Warthen

Bruce A. Kassebaum

Cover Designer

Team Coordinator

Team? We have a team?

Production

Special thanks to:

Charles R. Warthen, math and physics teacher, Renaissance Man, and lifelong friend;

Bruce A. Kassebaum, the best beta tester around, and lifelong friend;

Kathy Johnston, MD, Margaret Lambson, and Carol Wootton, for their support, encouragement, and for graciously providing test data.

Table of Contents

Foreword	0
Part I Introduction	5
1 Acknowledgments.....	5
2 What's in it for me?.....	6
Part II Run ADNA - an overview	7
Part III Menu Options	12
Part IV Controls and Shortcuts	14
Part V Processing data files	16
1 Select and edit the data file	19
2 Process the data file.....	20
3 Save or export the data.....	22
Part VI Using the exported data	22
Index	0

1 Introduction

ADNA is a utility for working with autosomal DNA data; specifically, for identifying overlaps in datasets prepared by Rob Warthen's excellent DNAGedCom.com site.

Update: Both Family Tree DNA and 23AndMe members can download their data directly from their respective web sites (September 2014).

ADNA reads the downloaded file into a grid, sorts the data by segment size in centiMorgans and discards any matches that do not meet the user-configurable threshold (set to 5 centiMorgans by default). It then iterates through the data, performing a brute-force comparison by chromosome and segment location, identifying overlapping segments.

Each name in the grid is arbitrarily assigned a color (pastel colors, since they are more easily perceived by color-deficient persons like me) so that all records for a specific person are rendered in the same color.

This application has been tested on several real-world datasets graciously provided to me for testing purposes by members of the genetic genealogy community; I am very grateful for their cooperation.

Although I'm not looking for another unpaid job, I am open to comments, constructive criticism, and bug reports at my [email](mailto:) address.

ADNA is written and compiled in Delphi XE7 from the current release of RAD Studio by Embarcadero Technologies; Delphi is an object-oriented Pascal for Windows.

Thank you for taking the time to try ADNA.

T. Wesley Erickson
Warner Springs, CA

1.1 Acknowledgments

Allow me to take a moment to explain how this project came about and to thank some of the people that made this application possible.

ADNA is the current iteration of an application written specifically to process autosomal DNA datasets prepared by Rob Warthen's excellent DNAGEDCOM.COM site.

[Charles R. Warthen](#)

I wrote ADNA because my lifelong friend, mentor, and high school physics teacher (46 years ago), Charlie Warthen, has embraced genetic genealogy. Charlie is a true Renaissance Man, who is equally at home discussing advanced mathematical algorithms, Renaissance musical instruments (many of which he plays), astronomy, photography, sailing, hiking, bird breeding... the list goes on for a long time.

When Charlie gets involved in an enterprise, he does so full measure, with no hesitation; and so it has been with DNA genealogy.

Margaret Lambson

Margaret is a steering committee member of the DNA Interest Group of the North San Diego County Genealogical Society, and she has been very helpful to newcomers like myself. She is always ready with insightful suggestions, feedback, and encouragement; she has also provided data for testing our various applications.

Carol Wootton

Carol has been particularly supportive during the pre-release stages of ADNA, providing data and serving as a beta tester.

Kathryn Johnston, MD

Kathy is an incredibly prolific and dedicated member of the genetic genealogy community; she serves as Administrator for several groups (including the SCGS DNA Project and the H2a2b Project) and moderates (with Bonny Cook) the DNA Interest Group of the Southern California Genealogical Society. Kathy is a frequent presenter on a variety of topics (most recently X-Chromosome), and has provided data and support for our applications.

Meanwhile, back at the ranch... The inspiration for ADNA can be traced back to two events: an informal discussion at an ISOGG Jamboree in Burbank, and a presentation on autosomal DNA by CeCe Moore to the DNA Interest Group of the North San Diego County Genealogical Society. Charlie and I had the same reaction to the description of the process of identifying overlaps that we did to the YDNA problem: here was a process that cried out for an automated solution; hence, this utility.

1.2 What's in it for me?

Who needs this program?

Actually, nobody. But if you are getting eyestrain or headaches from staring cross-eyed at row after row of unformatted numbers looking for overlapping ranges, you may find that it makes your life a little easier.

Why should I use it?

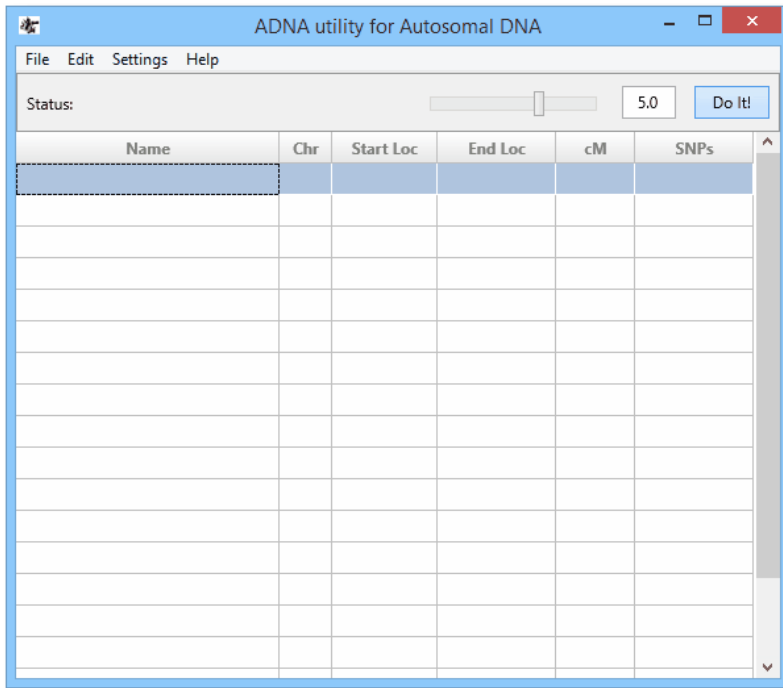
So you can spend your time in more enjoyable pursuits, like contacting people with overlapping segments to see how you might be related (You guys really know how to have a good time — I'll be on the balcony, sipping a good single-malt whiskey).

This utility may save time and prevent some common errors — assuming, of course, that it is working as intended...

Which is the other reason you should use it: to report bugs to [me](#) so that I can improve the program.

2 Run ADNA - an overview

The *Main* form is displayed on startup:



Brief nomenclature: ADNA's *Main* form consists of the following:

- Windows Menu (File, Edit, Settings and Help)
- Status label in header
- Slider to control color luminance
- Edit control to set threshold in centiMorgans
- **Do It!** button
- Grid to display data

A processed file yields a display similar to this:

Name	Ch	Start Loc	End Loc	cM	SNPs
Pam Bodwell	1	5118461	12148559	12.88	2085
Branda Wohlgenen	1	5118461	10595681	9.29	1596
Hayedeh Sepold Scott	1	19695400	26593874	9.24	2000
Carsten Edmund Malmberg	1	31908360	38625611	8.67	1665
Per-Olof Vilhelm Jonasson c/o Viktori	1	61885351	76790013	14.86	3700
Peter Ferdinand Reid	1	69857385	81219801	8.26	2600
James Lee Lowen	1	70453049	81219801	7.82	2500
Kjell Olof Yngve Yngvesson	1	89157869	101785313	11.1	3700
Scott Gregory Brookover	1	94382066	103926415	8.13	2800
James Spiegelberg	1	94382066	108398491	12.46	3800
Brenda Booth Decristina	1	94841509	106026345	9.39	3100
Amy B Siegle	1	95069962	105550721	8.65	2900
Susan C Griswold	1	96069014	115699858	19.82	5500
Peter Nelson	1	97081292	110902085	14.06	3800
Lonnie Gene Seawright	1	97459480	107100064	8.17	2500

Overlapping records are displayed in the *Matches* form:

Name	Chr	OL Start	OL Length	~cM
Brenda Booth Decristina	1	95,069,962	10,480,759	8.6
James Spiegelberg	1	95,069,962	10,480,759	8.6
Susan C Griswold	1	96,069,014	9,481,707	7.8
Scott Gregory Brookover	1	95,069,962	8,856,453	7.3
Peter Nelson	1	97,081,292	8,469,429	7.0
Lonnie Gene Seawright	1	97,459,480	8,091,241	6.7
Kjell Olof Yngve Yngvesson	1	95,069,962	6,715,351	5.5

The *Matches* form displays records that overlap the selected record on the left on the same chromosome

The user may elect to display all matches sorted either by Chromosome or Name:

The screenshot shows two windows from the ADNA software. The left window, titled 'Autosomal comparison results for Thomas Wesley Erickson', displays a table of 459 matches for 511 records, sorted by chromosome. The right window, titled 'Matches by Chromosome', displays the same data sorted by chromosome, with a search bar and 'Find', 'Copy', and 'Print' buttons.

Name	Chr	Start Loc	End Loc	cM	SNPs
Pam Bodwell	1	5118461	12148559	12.88	2085
Branda Wohlgenem	1	5118461	10595681	9.29	1596
Hayedeh Sepold Scott	1	19695400	26593874	9.24	2000
Carsten Edmund Malmberg	1	31908360	38625611	8.67	1665
Per-Olof Vilhelm Jonasson c/o Viktor	1	61885351	76790013	14.86	3700
Peter Ferdinand Reid	1	69857385	81219801	8.26	2600
James Lee Lowen	1	70453049	81219801	7.82	2500
Kjell Olof Yngve Yngvesson	1	89157869	101785313	11.1	3700
Scott Gregory Brookover	1	94382066	103926415	8.13	2800
James Spiegelberg	1	94382066	108398491	12.46	3800
Brenda Booth Decristina	1	94841509	106026345	9.39	3100
Amy B Siegle	1	95069962	105550721	8.65	2900
Susan C Griswold	1	96069014	115699858	19.82	5500
Peter Nelson	1	97081292	110902085	14.06	3800
Lonnie Gene Seawright	1	97459480	107100064	8.17	2500

Select **Output by Chromosome** to see all matching segments sorted by Chromosome

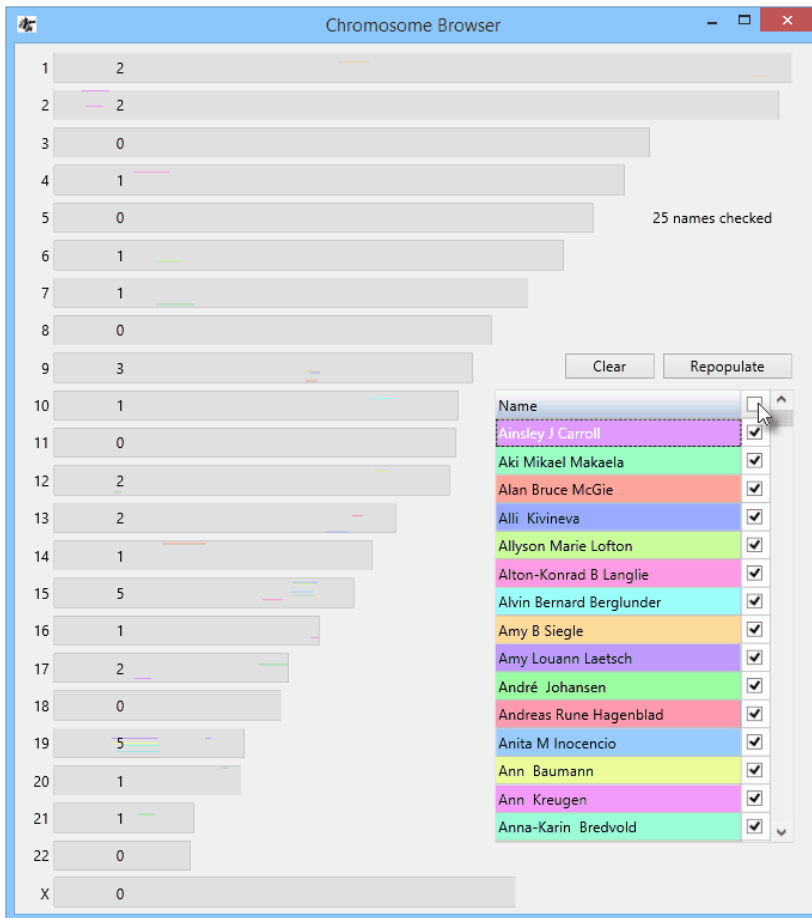
The screenshot shows two windows from the ADNA software. The left window, titled 'Autosomal comparison results for Thomas Wesley Erickson', displays a table of 459 matches for 511 records, sorted by name. The right window, titled 'Matches by Name', displays the same data sorted by name, with a search bar and 'Find', 'Copy', and 'Print' buttons.

Name	Chr	Start Loc	End Loc	cM	SNPs
Ainsley J Carroll	2	9044659	18449138	15.42	2682
Aki Mikael Makaan	20	55463163	58601464	8.61	1191
Alan Bruce McGie	14	36406412	50428544	10.25	2900
Alli Kivineva	15	79977687	88169071	12.26	2096
Allyson Marie Lofton	12	107280282	111796055	5.39	877
Allyson Marie Lofton	15	80231492	86213829	8.35	1396
Alton-Konrad B Langlie	4	26669873	38507991	12.58	2578
Alvin Bernard Berglunder	10	105378681	113943402	9.4	2300
Amy B Siegle	1	95069962	105550721	8.65	2900
Amy Louann Laetsch	19	19167147	34687643	8.51	1400
Amy Louann Laetsch	19	50499785	52637477	5.23	553
Andreas Rune Hagenblad	13	99682492	103173960	8.16	1400
André Johansen	17	68576138	78639702	25.68	3140
André Johansen	21	28127043	33448220	7.55	1397
Anita M Inocencio	15	79420887	86543964	9.79	1696

Select **Output by Name** to see all matching segments sorted by Name

Chromosome Browser

The Chromosome Browser is inspired by web applications that allow the user to graphically display segments for up to five people at a time. The "Name" list is populated with records selected by the user from the **Main** form (all records by default). Only the first 25 names are checked and displayed due to screen resolution limitations – fewer if less than 25 matching records exist. The number of segments mapped onto each chromosome appears towards the left.



Select Output to Browser to see selected segments on each chromosome

Click on the check box after each name to select or deselect that record. Select or deselect all records by clicking on the check box in the header.

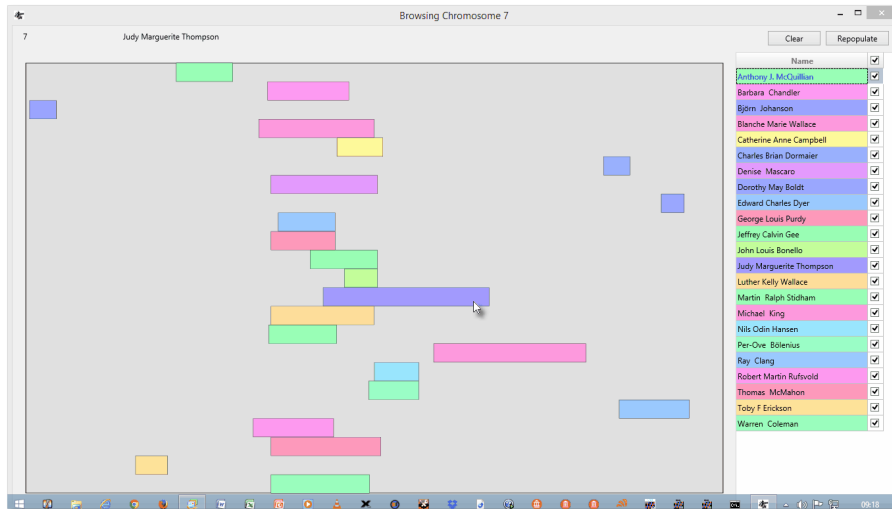
To display selected records, click on the **Repopulate** button. The number of currently selected names appears in a label above the name list.

To clear currently mapped segments, click on the **Clear** button.

Double-click on a chromosome to zoom in on a single chromosome (see **Detail** form).

Detail Form

Maps all selected segments (from Main form) for the selected chromosome:



Mapped segments for a single chromosome

Move the mouse cursor over a segment to display the owner (name) of that segment and its size in centiMorgans at the top of the screen.

Color Form

Use the color form to assign color to names (the same color will be assigned to all occurrences):

Name	Color
Ann Baumann	#FFEBFE9A
Ann Kreugen	#FFF59AFE
Anna-Karin Bredvold	#FF9AFED8
Annalisa Carlson	#FFFEBA9A
Anne Elizabeth Cleary	#FF9D9AFE
Anne Inger Marie Ölund	#FFB4FE9A
Anne Maria Ingman	#FFFE9AD1
Annick Hostetler	#FF9AEFFE
Annie Kathleen Nolan	#FFFEF19A
Anthony Fordham	#FFD39AFE
Anthony J. McQuillan	#FF9AFEB6
Arlene Folkers	#FFFE9A9B
Arne Wilhelm Horsten	#FF9AB8FE
B Sandvoll Larson	#FFD6FE9A
Barbara Chandler	#FFFE9AF3
Bengt Olovsson	#FF9AFEEC
Bengt Anders Auberg	#FFFE9AF3
Bengt Olof Engerstrom	#FFB19AFE
Bente Bernardsen	#FF9FFE9A

List of names with default colors

Name	Color
Ann Baumann	#FFEBFE9A
Ann Kreugen	#FFF59AFE
Anna-Karin Bredvold	#FF9AFED8
Annalisa Carlson	#FFFEBA9A
Anne Elizabeth Cleary	#FF9D9AFE
Anne Inger Marie Ölund	#FFB4FE9A
Anne Maria Ingman	#FFFE9AD1
Annick Hostetler	#FF9AEFFE
Annie Kathleen Nolan	#FFFEF19A
Anthony Fordham	#FFD39AFE
Anthony J. McQuillan	#FF9AFEB6
Arlene Folkers	#FFFE9A9B
Arne Wilhelm Horsten	#FF9AB8FE
B Sandvoll Larson	#FFD6FE9A
Barbara Chandler	#FFFE9AF3
Bengt Olovsson	#FF9AFEEC
Bengt Anders Auberg	#FFFE9AF3
Bengt Olof Engerstrom	#FFB19AFE
Bente Bernardsen	#FF9FFE9A

Select a color, click on "Apply" button

The Color form includes a (right-click) pop-up menu with an option to reset colors to default settings.

3 Menu Options

File | Open

Searches the user's Download folder for .csv (FTDNA) or .dat (23andMe) files. Both file types are displayed by default; to restrict display to one or the other, select a file extension from the file type drop-down list.

Family Tree DNA users: look for "nnnnnn_chromosome_browser_results.csv" where "nnnnnn" is the kit number.

23AndMe users: look for "23andMe_IBD_Segments_" followed by a name, date and time similar to this:

23andMe_IBD_Segments_Harold_Chancellor_20140920_120110.csv

A Brief Discussion of Legacy Files from DNAGEDCOM.com:

Legacy files from DNAGEDCOM.com are .zip files with the Kit Number (FTDNA) or email address (23AndMe) as the filename (e.g "37737.zip" or "name@site.com.zip"). The .zip file contains a folder named "files", which contains three files in comma-separated-values (.csv) format. The filenames end in "ChromosomeBrowser", "Family_Finder_Matches", and "ICW" (FTDNA) or "AC", "FIA", and "RF" (23AndMe). Use the "_ChromosomeBrowser.csv" file (FTDNA) or the "FIA.csv" file (23AndMe).

Key Point: Windows does not display filename extensions by default, so these files will likely appear as "nnnnn_ChromosomeBrowser" and "MyFile_FIA".

The "file type" drop-down also includes an option to display color profile (.cds) files, which allow the user to maintain multiple color profiles for a given data file.

File | Save...

Writes the contents of a processed file to a "comma-separated values" (.csv) file. This will typically create a file that is a subset of the original data file (with records below the Threshold discarded).

Note for Family Tree DNA users: this will overwrite a Family Tree DNA file of the same name; consider changing the filename before saving.

Note for 23AndMe users: the ".dat" file is in "comma-separated values" format; saving will create an additional file with the filename extension ".csv". Since Windows does not display filename extensions by default, the user will have two files with the same base filename. To avoid confusion, consider changing the filename before saving; we suggest appending the date in YYYYMMDD format.

File | Export

Exports the contents of processed file to a Microsoft Excel (.xls) file.

File | Exit

Closes the program.

Edit | Find Name

Searches the "Name" column for a name or partial name (use **F3** to search for successive matches).

Edit | Delete Rows

Deletes selected (highlighted) rows. Selected rows can also be deleted using the *Control-Delete* keystroke combination.

Edit | God Mode

Allows the user to drag and drop columns and otherwise mess around with the grid (warning: here be dragons). And yes, "God Mode" should really be under Settings.

Settings | Alternate Encoding

Switches between UTF-8 and Unicode character encoding. If diacritical marks do not appear correctly, change the setting of this menu item and reload the data file.

Settings | Delete non-matching

Discards records whose segments do not overlap other records. If not selected, non-matching records will remain in the table without a color assignment.

Settings | Output all by Chromosome

Displays the *Matches* form comprising a list of all matches for all selected records sorted by Chromosome and Start Location.

Settings | Output all by Name

Displays the *Matches* form comprising a list of all matches for all selected records sorted by Name.

Settings | Output to Browser

Displays selected records in a browser similar to the Chromosome Browser on the vendor's web sites. Double-click to zoom in on any chromosome.

If none of the above options is set, the *Matches* form is displayed with matches for the current (selected) record.

Settings | Open in Excel

Automatically opens a file in Excel (or other registered spreadsheet application) immediately after a File | Export operation.

Help | Contents

Displays this Help file.

Help | About

Displays the program "About" box.

4 Controls and Shortcuts

Main Form:

A status label in the header displays information about the number of records available or processed.

The following controls appear in the header of the main form:

Slider: controls the luminosity of the colors in the grid.

Threshold edit control: value in centiMorgans below which records are discarded during processing (5.0 by default).

Do It! button: processes the currently-open file.

Matches Form:

A status label in the header displays information about current matches.

The following controls appear in the header of the **Matches** form:

Find button: activates the local context-sensitive menu to search for names (including partial matches) in multi-record mode (user may also right-click).

Copy button: copies the data from the **Matches** form to the clipboard so that it can be pasted into another application (e.g. Excel).

Print button: dumps the contents of the **Matches** form to the default printer (no other print options currently available; let the author know if this is a priority for you).

Browser Form:

The *Browser* form is a graphical representation of segments mapped onto the various chromosomes.

A grid displays the names of all records currently selected on the main form; if the check box next to a name is checked, segments for that name are mapped on the display.

A status label indicates how many names are currently selected (maximum of 25 due to scaling limitations).

Detail Form:

The *Detail* form is a graphical mapping of all selected segments for a single chromosome. By default, all records in the Main Form table are selected, but the user has the ability to select a subset of records using any criteria (s)he prefers. The table can be sorted on any column, and multi-column sorts are supported by holding down the *Shift* key when selecting subsequent columns. This allows the user, for example, to sort the table by Chromosome and Start Location by clicking on the header of the Chromosome column, holding down the *Shift* key, then clicking on the "Start Loc" column. Clicking on any column a second time changes the sort order from ascending to descending; it is possible to mix sorts so one column is ascending and another is descending; feel free to experiment.

The user may select multiple contiguous or non-contiguous records for display using techniques previously described in this manual (standard Mac/Windows behavior).

The grid on the right side of the Detail form displays the names of all records currently selected from the main form. Records for which the check box is checked are mapped to the display; the current release is limited to 418 selected names on any chromosome at a time.

Color Form:

Allows the user to assign colors to names and to create multiple color profiles.

Accelerator Keys

ADNA follows the Windows convention of displaying accelerator keys with an underline when the "Alt" key is pressed; for example, to open a file, press *Alt-F*, *O*.

Function Keys

The following actions are assigned to function keys:

F1 : displays the Help file

F2 : activates Edit Mode in the grid (currently inactive)

F3 : finds successive matches following an **Edit | Find Name** command

F4 : in conjunction with the *Alt* key, closes the application

F5 : turns off display of segment (rectangle) borders

F6 : turns on display of segment (rectangle) borders

F8 : toggles "multi-select" mode in **Browser** and **Detail** forms

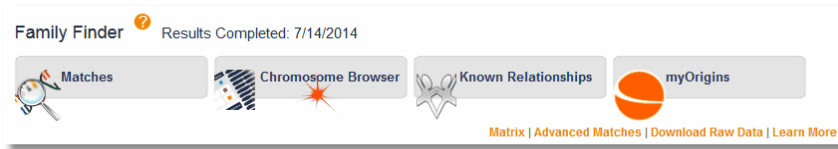
5 Processing data files

Actually, there are a couple of things that must be done first... (Side note: Did you ever notice that you can never do just one thing — you always have to do something else first?)

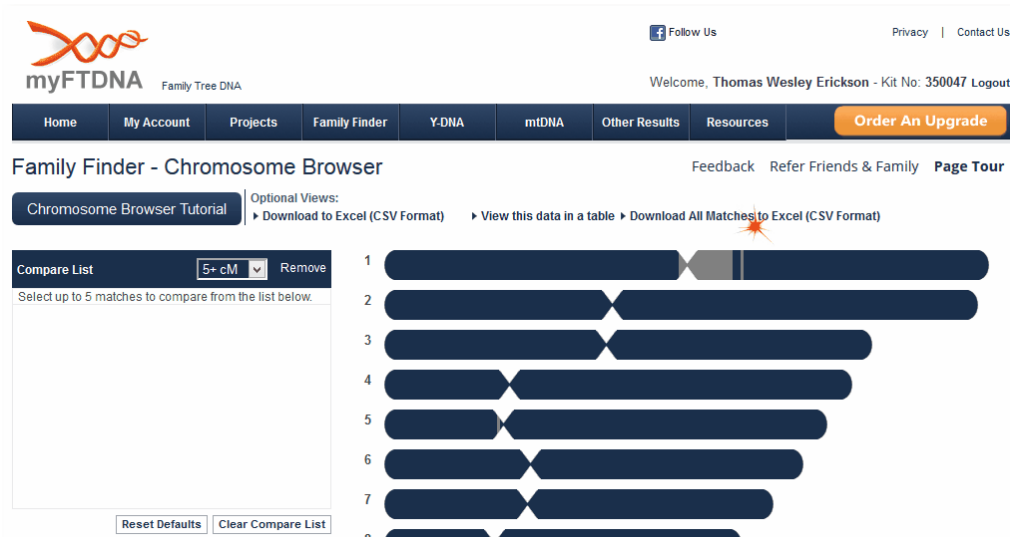
Since you are reading the Help File, I will conclude that you have successfully downloaded and installed ADNA (first steps complete!).

Family Tree DNA:

Log in to your FTDNA account and go to the [Chromosome Browser](#):

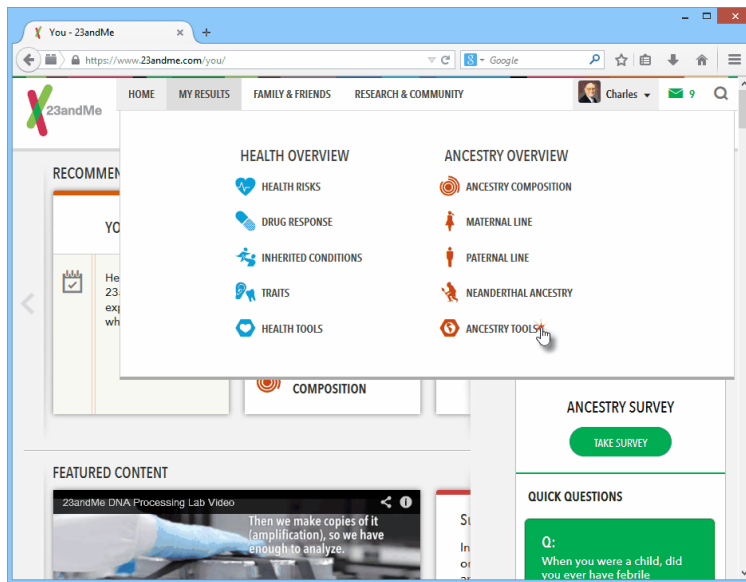


Click on [Download All Matches to Excel \(CSV Format\)](#) to generate a file named "nnnnnn_chromosome_browser_results.csv", where "nnnnnn" is the user's Kit Number):



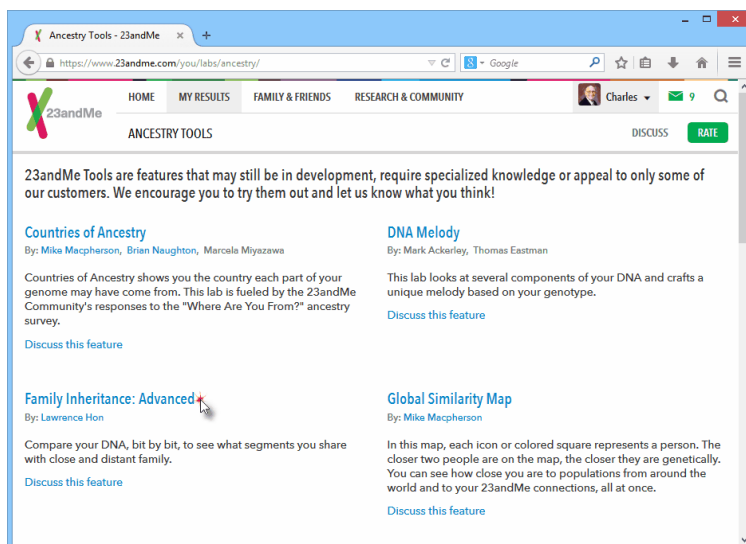
23AndMe:

Log into your 23AndMe account; from the Home Page, select **My Results | Ancestry Tools**:



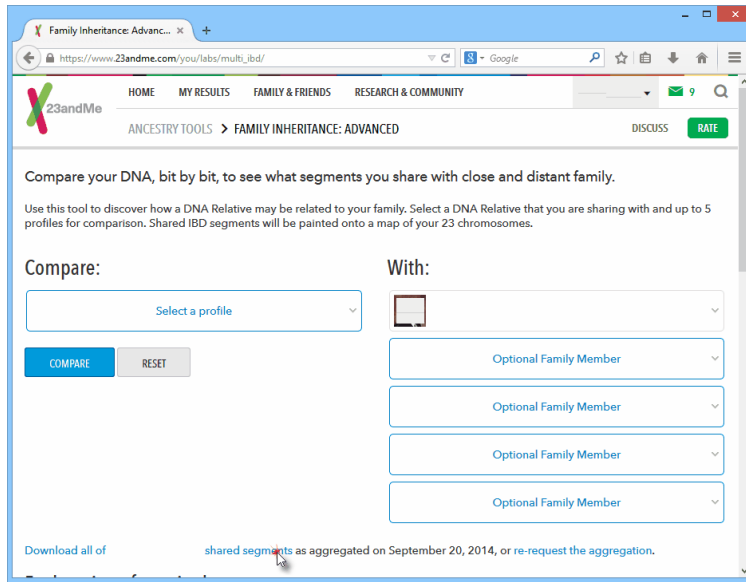
23AndMe: My Results | Ancestry Tools

From the **Ancestry Tools** page, select Family Inheritance: Advanced:



Select Family Inheritance Advanced

At the bottom of the screen click on the "Download all ... shared segments" link:



Click on the "Download all ... shared segments" link

You will receive an email with a link to download the assembled data file; click on the link to download the file to your computer.

The file will appear in the "Downloads" folder defined in your web browser; this usually corresponds to your personal "Downloads" folder.

If you use a download manager or have changed your web browser's settings, the file may download to another location; hopefully, you will know where to look for your files. If necessary, use the Windows "Search" function to locate your data file(s).

Extract the contents of the zip file to a known location; we suggest the "Downloads" folder.

Ancestry:

Currently unknown; if anyone can describe the process for downloading autosomal data, I will assess the viability of adding support for Ancestry to ADNA.

Okay, now we can get to work...

5.1 Select and edit the data file

Note: Family Tree DNA and 23AndMe members may now download their data directly as described in [Processing Data Files](#).

(The following Family Tree DNA information is deprecated)

Family Tree DNA: locate a file named "nnnnn_ChromosomeBrowser.csv" (where "nnnnn" is the Kit Number):

- nnnnn_ChromosomeBrowser.csv ←
- nnnnn_Family_Finder_Matches.csv
- nnnnn_ICW.csv

(The following 23AndMe information is deprecated)

23AndMe users: locate a file named "UserName-FIA.dat" (where "UserName" is your DNAGEDCOM.com user-id):

- UserName-AC.dat
- UserName-FIA.dat ←
- UserName-RF.dat

Key Point #1: Windows does not display filename extensions by default, so the ".csv" or ".dat" will likely not be visible.

Key Point #2: Zip files downloaded from DNAGEDCOM.com contain three files; ADNA uses only one of these. If you open one of the other files in ADNA, it will be identified as an "Invalid data file"; you will, however, be able to view the contents.

Select **File | Open** from the menu; navigate to and select a file, then click on the **Open** button to display the contents of the file in the grid.

Note: It may be desirable to remove records from the grid prior to processing; for example, you may want to remove known family members (e.g. parents or siblings) to prevent the display of large overlapping segments.

To delete a single row, select it (click on it), then:

- Press **Ctrl-Delete**, or
- Select **Edit | Delete Row(s)** from the menu

To delete multiple contiguous rows:

- Click on the first row, *shift-click* on the last, then press **Ctrl-Delete**
- Click on the first row; hold down *shift* key & press the *up-* or *down-arrow*, then press **Ctrl-Delete**

To delete multiple non-contiguous rows:

- Hold down the **Ctrl** key, click on multiple rows, then press **Ctrl-Delete**

The ADNA Utility does not have an "undo" feature at the present time (Sorry, Carol...).

To locate rows containing a name (or partial name), select **Edit | Find Name** from the menu, enter a name and click on the **OK** button. If a match is found, it will be selected and highlighted. To find subsequent matches, press the **F3** function key. If the **F3** function key does not work, focus may have moved away from the grid; click once anywhere on the grid then press **F3** again.

5.2 Process the data file

The default comparison threshold is 5.0 centiMorgans; to change this, enter a new value in the edit control. If the data file has already been processed, you will be prompted to reload the data file.

Check or un-check **Settings | Alternate** to affect the display of characters with diacritical marks.

Check **Settings | Delete non-matching** to delete non-matching records during processing.

Check **Settings | Output all by Chromosome** to display all matches sorted by Chromosome and Start Location.

Check **Settings | Output all by Name** to display all matches sorted by Name and Chromosome.

Check **Settings | Output to Browser** to graphically display segments for selected records

Check **Settings | Open in Excel** to immediately open a file after a **File | Export** operation.

Use the **File | Open** menu option to select a file for processing; if the user has not yet selected a file and clicks on the **Do It!** button, the file open dialog will be presented.

The **Do It!** button performs what it considers to be the next logical action based on context. For example, if the user has not opened a data file using the **File | Open** command and clicks on the **Do It!** button, the **File | Open** dialog will be displayed.

Do It! operations include the following:

- Display the File Open dialog to allow the user to select a data file.
- Process the currently-open file (sort the table, delete records whose segments are below the Threshold value in centiMorgans, identify overlapping records and assign a color to each name with matches). If **Delete non-matching** is checked, those records are deleted; if unchecked, they are displayed in the background color (white by default).
- Display the **Matches** form, the **Browser** form, or the **Detail** form (depending upon menu selection and current context).

Main Form

Once a data file has been processed, the user may review or edit the data (e.g. remove unneeded rows, sort the data for display purposes, etc.), or select rows to pass to one of the other forms.

Matches Form

Click on the **Find** button (or right-click on the grid) to search for a name. Select **Find Reference** (colored) or **Find Comparison** (matches) from the local menu. Enter a name or partial name and click on the **OK** button; the search is not case-sensitive. Press the **F3** function key to repeat the last search.

Click on the **Copy** button in the upper-right corner to copy the contents of the form to the clipboard for pasting into Excel or some other application.

Click on the **Print** button in the upper-right corner to send the contents of the form to the default printer.

Browser Form

Click on the **Clear** button to wipe the display clean for a new operation.

Click on the check boxes to select or deselect records for display; the current upper limit is 25 names. Click on the **All Records** check box in the header to select or deselect all rows.

Click on the **Repopulate** button to display the currently-selected records in the Names list.

Detail Form

Click on the **Clear** button to wipe the display clean for a new operation.

Click on the check boxes to select or deselect records for display; the current upper limit is 418 names. Click on the **All Records** check box in the header to select or deselect all rows.

Click on the **Repopulate** button to display all segments selected in the Main form for that chromosome.

To use Windows "multi-select" mode, first activate it by pressing the **F8** function key; a message will state that multi-select mode is active, and "Multi-Select" will appear in the **Names** grid header, then clear all currently-selected records by deselecting the **All Records** check box.

To select multiple contiguous records, click on the first record, then either:

- **Shift-Click** on the last record (hold down the **Shift** key and click); or
- Hold down the **Shift** key and press the **Up-Arrow** or **Down-Arrow** key to select multiple records.

To select multiple non-contiguous records:

- **Control-Click** on the desired records (hold the **Control** key down and click on all the records you want to select).

The ADNA Utility will facilitate the identification of persons with segments in common with the reference person. ADNA cannot determine if these segments are "Identical by Descent" or "Identical by State". The task of identifying candidates for follow-up to identify possible relatives is left to the user, but hopefully we have spared him or her some of the tedium.

5.3 Save or export the data

If you have edited your data file, you may want to save a copy to preserve those edits for future reference; for example, you might want to save a copy of the data with close family members removed (to minimize large overlapping segments). Note that, even if you delete rows, you will not be able to save a copy of the file until after the first processing pass (click the **Do It!** button).

Select **File | Save** from the Main Menu, assign a meaningful filename and click on the **Save** button to save a copy of the file in comma-separated values (.csv) format. Consider appending the date in YYYYMMDD format or a description of the changes made (e.g. MyFile-FIA-siblings-removed.csv).

Export to Excel (spreadsheet)

To save the processed data as an Excel (.xls) file, select **File | Export**, provide a filename, and click on the **Save** button.

If the **Settings | Open in Excel** menu item is checked, the processed data will be immediately opened in Excel (or other compatible spreadsheet application) following a **File | Save** operation.

6 Using the exported data

(Section reserved for subsequent update)



T. Wesley Erickson
PO Box 128
Warner Springs
CA 92086-0128