

THE PROGRAMMING CHALLENGES OF DOUBLE MATCH TRIANGULATOR

LOUIS KESSLER, DEVELOPER TALK

BYU FAMILY HISTORY TECHNOLOGY WORKSHOP, FEBRUARY 7, 2017

www.doublematchtriangulator.com

DOUBLE MATCH TRIANGULATOR IS ...

• An autosomal DNA analysis tool • A semi-finalist in the RootsTech 2017 Innovator Showdown • Conceptualized and developed by me starting Summer 2016 • Written in Delphi, which is essentially object extended Pascal • A Windows program that reads FamilyTreeDNA CBR files Creating an Excel output file to help analyze the DNA

CHROMOSOME BROWSER RESULTS FILE (CBR)

100000 Chromosome Browser Results 20160822.csv - Notepad П File Edit Format View Help Harry B00000, Joel S00001, 1, 170876383, 183564808, 10.64, 2800 Harry B00000, Joel S00001, 2, 1091127, 3733675, 5.26, 900 Harry B00000, Joel S00001, 2, 29300205, 33023313, 4.91, 1100 Harry B00000, Joel S00001, 2, 51848242, 54747547, 2.81, 900 Harry B00000, Joel S00001, 2, 72881715, 79080783, 7.33, 1600 Harry B00000, Joel S00001, 3, 36495, 4378767, 10.85, 1888 Harry B00000, Joel S00001, 4, 69117497, 71060137, 1.24, 500 Harry B00000, Joel S00001, 4, 114253251, 116767884, 1.97, 500 Harry B00000, Joel S00001, 5, 142643609, 145610930, 3.44, 700 Harry B00000, Joel S00001, 5, 167621874, 169059675, 2.62, 500 Harry B00000, Joel S00001, 5, 176763725, 180625733, 7.69, 1053 Harry B00000, Joel S00001, 6, 56101204, 63572255, 1.38, 500 Harry B00000, Joel S00001, 6, 72335742, 74438934, 1.56, 500 Harry B00000, Joel S00001, 7, 83468650, 86999456, 1.76, 700 Harry B00000, Joel S00001, 10, 2380581, 3387417, 2.74, 500 Harry B00000, Joel S00001, 10, 104404211, 124004929, 27.79, 5750 Harry B00000, Joel S00001, 11, 48014889, 57826280, 3.01, 1000 Harry B00000, Joel S00001, 12, 20668638, 21368250, 2.53, 500 Harry B00000, Joel S00001, 12, 33401350, 38671385, 1.26, 500 Harry B00000, Joel S00001, 12, 47590490, 50042536, 1.99, 500 Harry B00000, Joel S00001, 14, 61876003, 67118517, 4.59, 1100 Harry B00000, Joel S00001, 14, 99682846, 106358708, 9.35, 1198 Harry B00000, Joel S00001, 16, 74078162, 76241146, 4.45, 688 Harry B00000, Joel S00001, 17, 37316609, 39713299, 2.83, 500 Harry B00000, Joel S00001, 17, 71112486, 73584039, 5.93, 772 Harry B00000, Joel S00001, 18, 21019007, 23644650, 3.5, 1000 Harry B00000, Joel S00001, 18, 44772622, 46437311, 1.37, 500 Harry B00000, Joel S13869, 1, 34067461, 36938054, 3.91, 700 Harry B00000, Joel S13869, 1, 85420456, 87600621, 2.25, 597 Harry B00000, Joel \$13869, 1, 103128353, 107499183, 2.86, 800 Harry B00000, Joel S13869, 2, 57716395, 66550469, 7, 2000

<

×

DESIRED EXCEL OUTPUT

		_					-																															
	A VAME-A	B NAME-B	C NAME-C	D	-	F F	G	H CTADT DC		J	K MG-START D		M					_						_		_								HA HB				
			Beatrice I10789	_	START-AC			START-BC	163867390	3.35				Missing a-b Match	100 100	101	100 10	3 110	111 112	2 Irə	1/14 1/	5 110	100 1	10 113	100 1	51 102	103 104	+ 100 10	0 101	100 10	3 130	131 132	. 155	194 195	130 131	130 13	33 200	201
			Stephen F17636	-										Missing a-b Match				_				_		_					_		_						_	$ \longrightarrow $
				_	164755767		_		165765782				01.164.166		X a							_		_		_			_		_							i
			Laurence F14511		1 168113770				169516601	1.81	167284043 1			Missing a-b Match		ьγ	< X					_		_					_								_	$ \longrightarrow $
			Ofra W16059		1 168521569				172579436		167284043 1			Missing a-b Match Missing a-b Match				ь	ьь			_		_													_	$ \longrightarrow $
			Caren M11132 Marc H15014	-	1 167284043 1 168521569			168113770	169683826 169516601		167284043 1 167284043 1			Missing a-b Match		a 〉						_		_					_		_							$ \longrightarrow $
			Walter S18210		1 168521569			168113770			167284043 1					Ś		a a						_									+					
			Judith L14132		1 168521569	170011681		167869007			167284043 1			Missing a-b Match		ь Х		9						_														
			Samantha B17128		1 168521569			168113770			167284043 1			Missing a-b Match		b s		a						_									+					
			Rachel T16406		1 168521569			168113770						Missing a-b Match		Ś								_														
		Joel S00001			1 168521569			168113770						Missing a-b Match		Ś		_						_									+					
	,	Joel S00001			1 168521569			168113770						Missing a-b Match		>								_														
		Joel S00001			168521569			169008852						Missing a-b Match				ь						_														
204	Harry B00000	Joel S00001	Roger M16895	1	168113770	169516601	1.27	169008852	170011681					Missing a-b Match			X	ь						_									+					
		Joel S00001	-	1	1 169350022	170870563	1.24	167869007	169516601					Missing a-b Match		ьĿ	X	а																				
206	Harry B00000	Joel S00001	Marion C15123		1 170016288	173357037			173658364	4.57	167869007 2	00235705	01 167 200	Full Triangulation		b b	ь	X	X X	X		-								_	-						\rightarrow	1
		Joel S00001			1 170016288				171377576	3.67				Full Triangulation		ЬЕ	ь		Xa					_									+					
		Joel S00001			1 170876383				183564808	10.64	167869007 2							×	X X	Х	ΧХ	×	хх	×	х х	\times >	(<mark>.</mark>					<u> </u>	i. iii			in the second second		i. 👘
209	Harry B00000	Joel S00001	Sheldon S17409	1	167869007	179312044	8.85	171776209	178831136	5.08	167869007 2	00235705	01.167.200	Full Triangulation		a a) a	а	хх	х	х х	×	х х	а														í – T
210	Harry B00000	Joel S00001	Jeffrey S13624	1	167869007	179312044	8.85	171776209	177283106	4.16	167869007 2	00235705	01.167.200	Full Triangulation		a a	a a	а	х х	Х	х х	×	X a	а														1
211	Harry B00000	Joel S00001	devora g11886	1	1 171776209	176015986			176015986	2.99	167869007 2	00235705	01.167.200	Full Triangulation					ХХ	Х	х х	×																1
212	Harry B00000	Joel S00001	RACHEL S16402	1	170876383	173931568	2.07	171776209	174452682	1.63	167869007 2	00235705	01.167.200	Full Triangulation				а	х х	X	ь																	1
213	Harry B00000	Joel S00001	Richard B16551	1	169008852	200235705			183564808	9.66	167869007 2	00235705	01.167.200	Full Triangulation			а	а	a X	Х	ХХ	×	х х	×	х х	\times >	a	a a	a	a a	а	a a	a a	a a	a a	a a	а	1 1
			Daniel P11527		168521569			172592941		9.12	167869007 2	00235705	01.167.200	Full Triangulation		a	а а	а	a X	Х	х х	×	х х	×	х х	Xa	а	a a	а	a a	а	а						1 1
215	Harry B00000	Joel S00001	Adrienne M10122	1	1 171384394	181337647			181337647	8.03	167869007 2	00235705	01.167.200	Full Triangulation					a X	Х	х х	×	х х	×	X X													1 /
		Joel S00001			1 167869007			172592941						Full Triangulation		a a	a a	а	a X	х	х х	×	х х	×	×													1 /
			Padre N16087		169008852	180378611		172592941						Full Triangulation			а	а	a X	х	X X	×	х х	×	×													1
			Farrell H12495		1 167869007	178831136		172592941						Full Triangulation		a a	а	а	a X	х	X X	×	x x															i – 1
			Susana L17863		1 167869007				177723085					Full Triangulation		a a	а а	а	a X	X	××	×	×															i 1
			Phyllis R16333		1 171384394				177283106					Full Triangulation					a X	×	× ×	×	X a	а	a a	а			_		_							i – /
	/		Brian A11046		1 167869007				177723085					Full Triangulation		a a	за	а	a X	X	××	×	ь	_		_			_		_							i – /
			Natalie C15920		1 167869007				175552623					Full Triangulation		a a) а	а	a X	×	x x x x						,											i – /
			Valerie P18115		1 171776209 1 173661145		2.44	173357774	183011887 176015986	2.44				Full Triangulation				_	a a		XX	÷.	x x	×	x x	\times	a											i – 1
		Joel S00001	Julie G14190		1 173936525			173936525		2.44				Full Triangulation Full Triangulation				_		0	0 0	- C - I		×		XX											\rightarrow	i – 1
		Joel S00001			1 173936525			173936525						Full Triangulation						Ŷ	â â	- Â	ŝŝ	^	<u>^</u>	<u>^</u> (a	a a	а.	a a	a	a d	4			+	+	
			Mikhail A15676		1 176017112		2.09							Full Triangulation						0			ΩŶ										+	\rightarrow		++-		
	,		David W11714	-	1 176231846				178453336					Full Triangulation									χŶ										++	++		++-		
			Julie G14190		1 176231846				183564808					Full Triangulation						-		a	x x	×	x x	X X	a	a a	a	a a	a	a a	a			+		
	,		Phyllis R16333		171384394			177725444						Full Triangulation					a a	а	a a	a	X X	×	XX	×	-											
			Martin R15270		1 178061225			178461066						Full Triangulation								-	×	×	X X	XX	a	a a	a	a a	а	a a	a i	a a	a a	a a		
		Joel S00001			178834051				182502103					Full Triangulation									X	×	X X	X	а											
			Peter E16268		179312066			179312066						Full Triangulation										×	х х													
234	Harry B00000	Joel S00001	Nancy S15890	1	1 180786484	199655922	14.38	179653274	183564808					Full Triangulation										ь	х х	\times \times	a	a a	a	a a	а	a a	a i	a a	a a	a a		
235	Harry B00000	Joel S00001	Vladimir K18181	1	1 181339213	198335281	11.85	181339213	183564808					Full Triangulation											×	\times >	a	a a	a	a a	а	a a	a /	a a	a a	а		
236	Harry B00000	Joel S00001	sandra s17185	1	182508022	193221973	7.55	187869540	190685868	1.45	182508022 1	99655922	01.182.199	Missing a-b Match												a a	а	a a	X	хх	Х	a a	а					
237	Harry B00000	Joel S00001	Elaine A12150	_	183013710			184262239	186517433	1.97	182508022	186046132	01.182.186	Missing a-b Match												a	X	ХЬ									-	
			Laurie C14518		183568353				188874456					Missing a-b Match														XX	Ь	ь			+-+				+	
			Anita S10451					184814575			186533767 1			Missing a-b Match													Ь		X	a a	а		+		_			
			Fred L12593		1 188875431				190685868					Missing a-b Match												+ +				x x	-	a a	+	\rightarrow		+	+	
														,				-		-								-							_	+		\rightarrow

INPUT CHALLENGES

Handle input file problems
Merged by name
Non ASCII characters get mangled
Merge the input from 2 large files quickly
1. Load data first
2. Sequential compare

3. Sort results after

USER INTERFACE CHALLENGES

🞽 Double M	latch Triangulator 1.4	—		\times
File a:	100000_Chromosome_Browser_Results_20160822.csv v	· Harry B	00000	
○ File b:	100001_Chromosome_Browser_Results_20160608.csv D:\Documents\Personal\My Family Research\DNA\DMT Test Files\DMT Sample File	E CPD 6	0001 iles in folder	
	Include subfolders Analysis By Chromosome			
Press "Ru	n" to build 4 DMT files Rebuild 0 existing DMT fi	les		
Run	Cancel Folder a Help	FamilyTreeD	NA website	
				1
<				>

• KISS • One window • Input of: • Two files • File + Directory • Messages & Log File • Help File

EXCEL OUTPUT CHALLENGES

// See: http://stackoverflow.com/questions/39056055/office-automation-hangs-with-windows-CoInitializeEx(nil, COINIT_APARTMENTTHREADED or COINIT_DISABLE_OLE1DDE); oXL := CreateOleObject('Excel.Application'); oXL.Visible := False; oXL.ScreenUpdating := False; { From: http://stackoverflow.com/questions/1436737 oXL.DisplayAlerts := false;

oXL.Workbooks.open(MapFilename, 1 { ReadOnly });

oSheet := oXL.ActiveSheet; if ChromosomeNum = 0 then

oSheet.Name := 'Map'

else

oSheet.Name := 'Map ' + inttostr(ChromosomeNum);

```
if ChromosomeNum = 0 then { Move the People Info to its own page } begin
    o2Sheet := oXL.Sheets.Add(NULL,oSheet,NULL,NULL);
    o2Sheet.Name := 'People';
    oXL.ActiveWindow.Zoom := 70;
```

```
oRng := oSheet.Range['A' + inttostr(MatchAnalysis.Count + 2) + ':AK' + inttostr(N
oRng.Cut;
o2Sheet.Select;
oXL.ActiveSheet.Paste;
oSheet.Select;
end;
```

oSheet.Range['A1:N1'].ColumnWidth := 0.1; oSheet.Range['02:02'].Select; oXL.ActiveWindow.FreezePanes := True; oXL.ActiveWindow.Zoom := 70;

Excel Automation is NOT well Documented in its Delphi form

EXCEL OUTPUT CHALLENGES (CONTINUED)

- Requires its own thread (or you'll get non-response)
 Cannot cancel it if it fails
- Uses Variant Arrays
 - Using variant arrays to fill 200,000 rows is sloooow.
 - Create CSV file takes 0.5 seconds. Create Excel file 100 seconds.
 - Solution found: Create CSV file and let Excel load it 3 seconds.
- Requires Excel to be installed

DOUBLE MATCH TRIANGULATOR

Developer Questions?