JIT HEALTH SCIENCE CENTER



Analysis of Massively Parallel Sequencing Data Using the STR Allele Identification Tool – Razor (STRait Razor)

Frank Wendt, Ph.D. Candidate

Institute of Applied Genetics
Department of Molecular and Medical Genetics
University of North Texas Health Science Center
3500 Camp Bowie Blvd.
Fort Worth, Texas 76107

Frank.Wendt@my.unthsc.edu

Lab Phone: 817-765-2940

Overview

- How STRait Razor works
- Starting your analysis
- Analysis workbook
- Customization potential
- Try it out



Overview

- How STRait Razor works
- Starting your analysis
- Analysis workbook
- Customization potential
- Try it out





4							
CSF1PO	AUTOSOMAL	AAGATAGATAGATT	AGGAAGTACTTA	TAAGTACTTCCT	AATCTATCTATCTT	AGATAGATAGAT ATCTATCTATCT	CSF1PO:5=20;CSF1PO:5.1=21;CSF1PO:5.2=22;CSF1
D10S1248	AUTOSOMAL	TATTGTCTTCAT	ACTCACTCATTT	AAATGAGTGAGT	ATGAAGACAATA	TTCCTTCC GGAAGGAAGGAA	D10S1248:8=32;D10S1248:8.1=33;D10S1248:8.2=34;
D12S391	AUTOSOMAL	AAATCCCCTCTC	ACCTATGCATCC	GGATGCATAGGT	GAGAGGGGATTT	GTCTGTCTGTCT AGACAGACAGAC	D12S391:15=60;D12S391:15.1=61;D12S391:15.2=62;



Locus

	1						
CSF1PO	AUTOSOMAL	AAGATAGATAGATT	AGGAAGTACTTA	TAAGTACTTCCT	AATCTATCTATCTT	AGATAGATAGAT ATCTATCTATCT	CSF1PO:5=20;CSF1PO:5.1=21;CSF1PO:5.2=22;CSF1
D10S1248	AUTOSOMAL	TATTGTCTTCAT	ACTCACTCATTT	AAATGAGTGAGT	ATGAAGACAATA	TTCCTTCC GGAAGGAAGGAA	D10S1248:8=32;D10S1248:8.1=33;D10S1248:8.2=34;
D12S391	AUTOSOMAL	AAATCCCCTCTC	ACCTATGCATCC	GGATGCATAGGT	GAGAGGGGATTT	GTCTGTCTGTCT AGACAGACAGAC	D12S391:15=60;D12S391:15.1=61;D12S391:15.2=62;



- Locus
- Type of locus

CSF1PO	AUTOSOMAL	AAGATAGATAGATT	AGGAAGTACTTA	TAAGTACTTCCT	AATCTATCTATCTT	AGATAGATAGAT ATCTATCT	CSF1PO:5=20;CSF1PO:5.1=21;CSF1PO:5.2=22;CSF1P
D10S1248	AUTOSOMAL	TATTGTCTTCAT	ACTCACTCATTT	AAATGAGTGAGT	ATGAAGACAATA	TTCCTTCC GGAAGGAAGGAA	D10S1248:8=32;D10S1248:8.1=33;D10S1248:8.2=34;
D12S391	AUTOSOMAL	AAATCCCCTCTC	ACCTATGCATCC	GGATGCATAGGT	GAGAGGGGATTT	GTCTGTCTGTCT AGACAGACAGAC	D12S391:15=60;D12S391:15.1=61;D12S391:15.2=62;



- Locus
- Type of locus
- Flanking region sequences

CSF1PO	AUTOSOMAL	AAGATAGATAGATT	AGGAAGTACTTA	TAAGTACTTCCT	AATCTATCTATCTT	AGATAGATAGAT ATCTATCT	CSF1PO:5=20;CSF1PO:5.1=21;CSF1PO:5.2=22;CSF1
D10S1248	AUTOSOMAL	TATTGTCTTCAT	ACTCACTCATTT	AAATGAGTGAGT	ATGAAGACAATA	TTCCTTCCTTCC GGAAGGAAGGAA	D10S1248:8=32;D10S1248:8.1=33;D10S1248:8.2=34;
D12S391	AUTOSOMAL	AAATCCCCTCTC	ACCTATGCATCC	GGATGCATAGGT	GAGAGGGGATTT	GTCTGTCTGTCT AGACAGACAGAC	D12S391:15=60;D12S391:15.1=61;D12S391:15.2=62
		ll .					



- Locus
- Type of locus
- Flanking region sequences
- Unique sequence between the flanking regions

CSF1PO	AUTOSOMAL	AAGATAGATAGATT	AGGAAGTACTTA	TAAGTACTTCCT	AATCTATCTATCTT	AGATAGATAGAT ATCTATCT	CSF1PO:5=20;CSF1PO:5.1=21;CSF1PO:5.2=22;CSF1
D10S1248	AUTOSOMAL	TATTGTCTTCAT	ACTCACTCATTT	AAATGAGTGAGT	ATGAAGACAATA	TTCCTTCCTTCC GGAAGGAAGGAA	D10S1248:8=32;D10S1248:8.1=33;D10S1248:8.2=34;
D12S391	AUTOSOMAL	AAATCCCCTCTC	ACCTATGCATCC	GGATGCATAGGT	GAGAGGGGATTT	GTCTGTCTGTCT AGACAGACAGAC	D12S391:15=60;D12S391:15.1=61;D12S391:15.2=62



- Locus
- Type of locus
- Flanking region sequences
- Unique sequence between the flanking regions
- Allele designation based on length between flanks

AUTOSOMAL	AAGATAGATAGATT	AGGAAGTACTTA	TAAGTACTTCCT	AATCTATCTATCTT	AGATAGATAGAT ATCTATCT	CSF1PO:5=20;CSF1PO:5.1=21;CSF1PO:5.2=22;CSF1
AUTOSOMAL	TATTGTCTTCAT	ACTCACTCATTT	AAATGAGTGAGT	ATGAAGACAATA	TTCCTTCCTTCC GGAAGGAAGGAA	D10S1248:8=32;D10S1248:8.1=33;D10S1248:8.2=34
AUTOSOMAL	AAATCCCCTCTC	ACCTATGCATCC	GGATGCATAGGT	GAGAGGGGATTT	GTCTGTCTGTCT AGACAGACAGAC	D12S391:15=60;D12S391:15.1=61;D12S391:15.2=62
4	AUTOSOMAL	AUTOSOMAL AAGATAGATAGATT AUTOSOMAL TATTGTCTTCAT AUTOSOMAL AAATCCCCTCTC	AUTOSOMAL TATTGTCTTCAT ACTCACTCATTT	AUTOSOMAL TATTGTCTTCAT ACTCACTCATTT AAATGAGTGAGT	AUTOSOMAL TATTGTCTTCAT ACTCACTCATTT AAATGAGTGAGT ATGAAGACAATA	AUTOSOMAL TATTGTCTTCAT ACTCACTCATTT AAATGAGTGAGT ATGAAGACAATA TTCCTTCCTGGAAGGAAGGAA



Overview

- How STRait Razor works
- Starting your analysis
- Analysis workbook
- Customization potential
- Try it out



Download File

- https://www.unthsc.edu/graduate-school-of-biomedicalsciences/molecular-and-medical-genetics/laboratory-facultyand-staff/strait-razor/
- Unzip the downloaded STRait Razor v2.5 zipped file

Important contents:

Name	Date modified	Туре	Size
forenseq	8/29/2016 11:01 AM	CONFIG File	52 KB
forenseqbeta	8/29/2016 11:01 AM	CONFIG File	58 KB
fusion fusion	8/29/2016 11:01 AM	CONFIG File	24 KB
globalfiler	8/29/2016 11:01 AM	CONFIG File	23 KB
locus	8/29/2016 11:01 AM	CONFIG File	86 KB
PGM25	8/29/2016 11:01 AM	CONFIG File	21 KB
ppss	8/29/2016 11:01 AM	File	83 KB
readfq.pl	8/29/2016 11:01 AM	PL File	2 KB
readme .	8/29/2016 11:01 AM	TXT File	3 KB
runSTRait.sh	8/29/2016 11:01 AM	SH File	6 KB
STRaitRazor.pl	8/29/2016 11:01 AM	PL File	7 KB
STRaitRazorDB.pl	8/29/2016 11:01 AM	PL File	5 KB
■ tre-0.8.0.tar	8/29/2016 11:01 AM	GZ File	471 KB



Download File

- https://www.unthsc.edu/graduate-school-of-biomedicalsciences/molecular-and-medical-genetics/laboratory-facultyand-staff/strait-razor/
- Unzip the downloaded STRait Razor v2.5 zipped file

Important contents:

Name	Date modified	Туре	Size
forenseq	8/29/2016 11:01 AM	CONFIG File	52 KB
forenseqbeta	8/29/2016 11:01 AM	CONFIG File	58 KB
fusion fusion	8/29/2016 11:01 AM	CONFIG File	24 KB
globalfiler	8/29/2016 11:01 AM	CONFIG File	23 KB
locus	8/29/2016 11:01 AM	CONFIG File	86 KB
PGM25	8/29/2016 11:01 AM	CONFIG File	21 KB
] ppss	8/29/2016 11:01 AM	File	83 KB
readfq.pl	8/29/2016 11:01 AM	PL File	2 KB
readme readme	8/29/2016 11:01 AM	TXT File	3 KB
runSTRait.sh	8/29/2016 11:01 AM	SH File	6 KB
STRaitRazor.pl	8/29/2016 11:01 AM	PL File	7 KB
STRaitRazorDB.pl	8/29/2016 11:01 AM	PL File	5 KB
™ tre-0.8.0.tar	8/29/2016 11:01 AM	GZ File	471 KB



Creating Shell Script

Open "runSTRait.sh" file in a text editor

```
#!/bin/bash
//STRaitRazor.pl -dir /illumina/data/FGx/Analyses -fastq /illumina/data/FGx/Analyses/1 - Copy (13)/R701-A501_S1_L001_R1_001.fastq -SampleNum R701-A501_S1_L001_R1_001 -typeselection ALL
//STRaitRazor.pl -dir /illumina/data/FGx/Analyses -fastq /illumina/data/FGx/Analyses/1 - Copy (13)/R701-A502_S2_L001_R1_001.fastq -SampleNum R701-A502_S2_L001_R1_001 -typeselection ALL
//STRaitRazor.pl -dir /illumina/data/FGx/Analyses -fastq /illumina/data/FGx/Analyses -fas
```

```
| STRaitRazor.pl -dir /illumina/data/FRW/straitrazorresults -fastq /illumina/data/FRW/indels/011 SR L001 R1 001.fastq -fastq /illumina/data/FRW/indels/021 S9 L001 R2 001.fastq.fastq -SampleNum 021 S9 L001 -typeselection ALL -locusconfig indellocus.config /STRaitRazor.pl -dir /illumina/data/FRW/straitrazorresults -fastq /illumina/data/FRW/indels/022 S10 L001 R2 001.fastq.fastq -SampleNum 021 S9 L001 -typeselection ALL -locusconfig indellocus.config /STRaitRazor.pl -dir /illumina/data/FRW/straitrazorresults -fastq /illumina/data/FRW/indels/022 S10 L001 R2 001.fastq.fastq -SampleNum 022 S10 L001 -typeselection ALL -locusconfig indellocus.config /STRaitRazor.pl -dir /illumina/data/FRW/straitrazorresults -fastq /illumina/data/FRW/indels/023 S11 L001 R2 001.fastq -fastq -fastq -fastq /illumina/data/FRW/indels/023 S11 L001 R2 001.fastq -fastq -fastq /illumina/data/FRW/indels/023 S01 L001 -typeselection ALL -locusconfig indellocus.config /Illumina/data/FRW/straitrazorresults -fastq /illumina/data/FRW/indels/023 S01 L001 -typeselection ALL -locusconfig indellocus.config /Illumina/data/FRW/straitraz
```



Executing Shell Script

```
Welcome to STRait Razor -- the STR allele identification tool!
Extracting sequences from FASTQ file(s)
Now extracting reads encompassing the repeat sequences (including flanking
regions with mismatches up to the user-defined limits).
Trimming will also be performed. This process will be performed in parallel...
Aug 30 06:34:33:
Aug 30 06:34:33: ======
Aug 30 06:34:33:
                                        |P|P|S|S|
Aug 30 06:34:33: Distributed Parallel Processing Shell Script vers. 2.85
Aug 30 06:34:33: ===
Aug 30 06:34:33: Hostname:
                                       ic-tthsc
Aug 30 06:34:33: -----
Aug 30 06:34:33: CPU: Intel(R) Xeon(R) CPU E7430 @ 2.13GHz
Aug 30 06:34:33: Found 16 logic processors.
Aug 30 06:34:33: Starting 16 parallel workers.
Aug 30 06:34:33: -----
Aug 30 06:38:37: Currently 53 percent complete. Processed 32 of 60. ☐
```



Executing Shell Script

```
Welcome to STRait Razor -- the STR allele identification tool!
Extracting sequences from FASTQ file(s)
Now extracting reads encompassing the repeat sequences (including flanking
regions with mismatches up to the user-defined limits).
Trimming will also be performed. This process will be performed in parallel...
Aug 30 06:17:36:
Aug 30 06:17:36:
Aug 30 06:17:36:
                                     |P|P|S|S|
Aug 30 06:17:36: Distributed Parallel Processing Shell Script vers. 2.85
Aug 30 06:17:36: Hostname:
                                    ic-tthsc
Aug 30 06:17:36: -----
Aug 30 06:17:36: CPU: Intel(R) Xeon(R) CPU E7430 @ 2.13GHz
Aug 30 06:17:36: Found 16 logic processors.
Aug 30 06:17:36: Starting 16 parallel workers.
Aug 30 06:17:36: -----
Aug 30 06:34:00: One job is remaining.
Aug 30 06:34:25: Total processing time (hh:mm:ss): 00:16:49
Aug 30 06:34:25: Finished. Consult ppss dir/job log for job output.
Extraction complete. Flanks and surrounding sequences have also been trimmed
from repeat sequences. Now writing allele calls to text file ...
Computation complete.
```



Output

↓ 4A_S7_L001	1/14/2016 12:13 PM	File folder	
005A_S5_L001	1/14/2016 10:20 AM	File folder	
№ 007A_S6_L001	1/14/2016 11:05 AM	File folder	
	I.		
<mark>}} R1</mark>	1/14/2016 12:13 PM	File folder	
<u></u> R2 R2	1/14/2016 12:13 PM	File folder	
rawSTRcallsR1	1/14/2016 12:13 PM	TXT File	6 KE
rawSTRcallsR2	1/14/2016 12:13 PM	TXT File	6 KE
allsequences	1/14/2016 11:55 AM	TXT File	12 KB
rs4187.Allelecalls	1/14/2016 11:35 AM	TXT File	1 KB
rs4187	1/14/2016 11:35 AM	SEQUENCES File	1 KB
rs16402.Allelecalls	1/14/2016 11:35 AM	TXT File	1 KB
rs16402	1/14/2016 11:35 AM	SEQUENCES File	1 KB
rs16458.Allelecalls	1/14/2016 11:50 AM	TXT File	1 KB
z rs16458	1/14/2016 11:50 AM	SEQUENCES File	1 KB
rs16624.Allelecalls	1/14/2016 11:35 AM	TXT File	1 KB
☑ rs16624	1/14/2016 11:35 AM	SEQUENCES File	1 KB



Output

↓ 4A_S7_L001	1/14/2016 12:13 PM	File folder	
₩ 005A_S5_L001	1/14/2016 10:20 AM	File folder	
₩ 007A_S6_L001	1/14/2016 11:05 AM	File folder	
<mark>}}</mark>	1/14/2016 12:13 PM	File folder	
<u></u>	1/14/2016 12:13 PM	File folder	
rawSTRcallsR1	1/14/2016 12:13 PM	TXT File	6 KB
rawSTRcallsR2	1/14/2016 12:13 PM	TXT File	6 KB
	•		
allsequences	1/14/2016 11:55 AM	TXT File	12 KB
allsequences rs4187.Allelecalls	1/14/2016 11:55 AM 1/14/2016 11:35 AM	TXT File TXT File	
allsequences rs4187.Allelecalls rs4187	1/14/2016 11:55 AM 1/14/2016 11:35 AM 1/14/2016 11:35 AM	TXT File TXT File SEQUENCES File	12 KB 1 KB 1 KB
rs4187.Allelecalls	1/14/2016 11:35 AM	TXT File	1 KB
rs4187.Allelecalls rs4187	1/14/2016 11:35 AM 1/14/2016 11:35 AM	TXT File SEQUENCES File	1 KB 1 KB
rs4187.Allelecalls rs4187 rs16402.Allelecalls	1/14/2016 11:35 AM 1/14/2016 11:35 AM 1/14/2016 11:35 AM	TXT File SEQUENCES File TXT File	1 KB 1 KB 1 KB
rs4187.Allelecalls rs4187 rs16402.Allelecalls rs16402	1/14/2016 11:35 AM 1/14/2016 11:35 AM 1/14/2016 11:35 AM 1/14/2016 11:35 AM	TXT File SEQUENCES File TXT File SEQUENCES File	1 KB 1 KB 1 KB 1 KB
rs4187.Allelecalls rs4187 rs16402.Allelecalls rs16402 rs16458.Allelecalls	1/14/2016 11:35 AM 1/14/2016 11:35 AM 1/14/2016 11:35 AM 1/14/2016 11:35 AM 1/14/2016 11:50 AM	TXT File SEQUENCES File TXT File SEQUENCES File TXT File	1 KB 1 KB 1 KB 1 KB



Overview

- How STRait Razor works
- Starting your analysis
- Analysis workbook
- Customization potential
- Try it out



Open Analysis Workbook

Include sample	Sample Name		Note: Compile Sample List button allows the user to		Sample List	
¥	011_S8_L001		select the folder containing the STRait Razor Results.	1	011_S8_L001	Z:\FRW\NewINDEL\011_S8_L001\R1\allsequences.txt
~	011_S8_L001	Compile	From there, you can remove unwanted samples	2	011_S8_L001	Z:\FRW\NewINDEL\011_S8_L001\R2\allsequences.txt
~	021_S9_L001	Sample	simply by removing the "check" or "a" from column C	3	021_S9_L001	Z:\FRW\NewINDEL\021_S9_L001\R1\allsequences.txt
J	021_S9_L001	List		4	021_S9_L001	Z:\FRW\NewINDEL\021_S9_L001\R2\allsequences.txt
~	022_S10_L001			5	022_S10_L001	Z:\FRW\NewINDEL\022_S10_L001\R1\allsequences.txt
v	022_S10_L001	Process		6	022_S10_L001	Z:\FRW\NewINDEL\022_S10_L001\R2\allsequences.txt
-	023_S11_L001	Queue		7	023_S11_L001	Z:\FRW\NewINDEL\023_S11_L001\R1\allsequences.txt
~	023_S11_L001			8	023_S11_L001	Z:\FRW\NewINDEL\023_S11_L001\R2\allsequences.txt
~	00029_S11_L001			9	00029_S11_L001	Z:\FRW\NewINDEL\00029_S11_L001\R1\allsequences.txt
J	00029_S11_L001	# of Reads		10	00029_S11_L001	Z:\FRW\NewINDEL\00029_S11_L001\R2\allsequences.txt
~	47A_S9_L001	Paired End		11	47A_S9_L001	Z:\FRW\NewINDEL\47A_S9_L001\R1\allsequences.txt
v	47A_S9_L001			12	47A_S9_L001	Z:\FRW\NewINDEL\47A_S9_L001\R2\allsequences.txt
	51_S5_L001			13	51_S5_L001	Z:\FRW\NewINDEL\51_S5_L001\R1\allsequences.txt
~	51_S5_L001			14	51_S5_L001	Z:\FRW\NewINDEL\51_S5_L001\R2\allsequences.txt
~	59_S6_L001			15	59_S6_L001	Z:\FRW\NewINDEL\59_S6_L001\R1\allsequences.txt
~	59_S6_L001			16	59_S6_L001	Z:\FRW\NewINDEL\59_S6_L001\R2\allsequences.txt
~	74_S7_L001			17	74_S7_L001	Z:\FRW\NewINDEL\74_S7_L001\R1\allsequences.txt
~	74_S7_L001			18	74_S7_L001	Z:\FRW\NewINDEL\74_S7_L001\R2\allsequences.txt

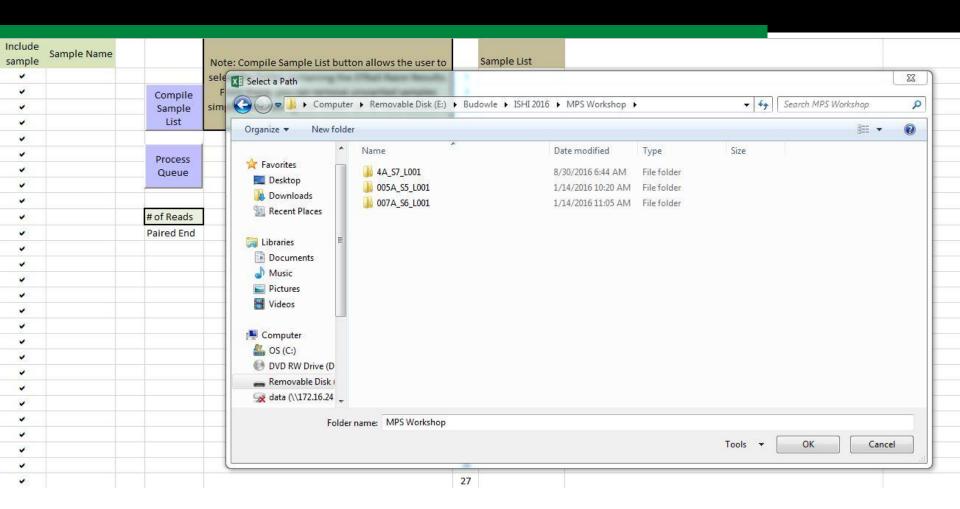


Open Analysis Workbook

Include sample	Sample Name		Note: Compile Sample List button allows the user to		Sample List	
~	011_S8_L001		elect the folder containing the STRait Razor Results.	1	011_S8_L001	Z:\FRW\NewINDEL\011_S8_L001\R1\allsequences.txt
~	011_S8_L001	Compile	From there, you can remove unwanted samples	2	011_S8_L001	Z:\FRW\NewINDEL\011_S8_L001\R2\allsequences.txt
~	021_S9_L001	Sample	imply by removing the "check" or "a" from column C	3	021_S9_L001	Z:\FRW\NewINDEL\021_S9_L001\R1\allsequences.txt
J	021_S9_L001	List		4	021_S9_L001	Z:\FRW\NewINDEL\021_S9_L001\R2\allsequences.txt
~	022_S10_L001			5	022_S10_L001	Z:\FRW\NewINDEL\022_S10_L001\R1\allsequences.txt
v	022_S10_L001	Process		6	022_S10_L001	Z:\FRW\NewINDEL\022_S10_L001\R2\allsequences.txt
-	023_S11_L001	Queue		7	023_S11_L001	Z:\FRW\NewINDEL\023_S11_L001\R1\allsequences.txt
~	023_S11_L001			8	023_S11_L001	Z:\FRW\NewINDEL\023_S11_L001\R2\allsequences.txt
V	00029_S11_L001			9	00029_S11_L001	Z:\FRW\NewINDEL\00029_S11_L001\R1\allsequences.txt
J	00029_S11_L001	# of Reads		10	00029_S11_L001	Z:\FRW\NewINDEL\00029_S11_L001\R2\allsequences.txt
~	47A_S9_L001	Paired End		11	47A_S9_L001	Z:\FRW\NewINDEL\47A_S9_L001\R1\allsequences.txt
~	47A_S9_L001			12	47A_S9_L001	Z:\FRW\NewINDEL\47A_S9_L001\R2\allsequences.txt
J	51_S5_L001			13	51_S5_L001	Z:\FRW\NewINDEL\51_S5_L001\R1\allsequences.txt
~	51_S5_L001			14	51_S5_L001	Z:\FRW\NewINDEL\51_S5_L001\R2\allsequences.txt
~	59_S6_L001			15	59_S6_L001	Z:\FRW\NewINDEL\59_S6_L001\R1\allsequences.txt
~	59_S6_L001			16	59_S6_L001	Z:\FRW\NewINDEL\59_S6_L001\R2\allsequences.txt
~	74_S7_L001			17	74_S7_L001	Z:\FRW\NewINDEL\74_S7_L001\R1\allsequences.txt
~	74_S7_L001			18	74 S7 L001	Z:\FRW\NewINDEL\74_S7_L001\R2\allsequences.txt



Navigate to STRait Razor Output





Load allsequences.txt

Include sample	Sample Name		Note: Compile Sample List button allows the user to		Sample List	
~	007A_S6_L001		select the folder containing the STRait Razor Results.	1	007A_S6_L001	E:\Budowle\ISHI 2016\MPS Workshop\007A_S6_L001\R1\allsequences.txt
~	007A_S6_L001	Compile	From there, you can remove unwanted samples	2	007A_S6_L001	E:\Budowle\ISHI 2016\MPS Workshop\007A_S6_L001\R2\allsequences.txt
,	4A_S7_L001	Sample	simply by removing the "check" or "a" from column C	3	4A_S7_L001	E:\Budowle\ISHI 2016\MPS Workshop\4A_S7_L001\R1\allsequences.txt
~	4A_S7_L001	List		4	4A_S7_L001	E:\Budowle\ISHI 2016\MPS Workshop\4A_S7_L001\R2\allsequences.txt
v	005A_S5_L001			5	005A_S5_L001	E:\Budowle\ISHI 2016\MPS Workshop\005A_S5_L001\R1\allsequences.txt
~	005A_S5_L001	Process		6	005A_S5_L001	E:\Budowle\ISHI 2016\MPS Workshop\005A_S5_L001\R2\allsequences.txt
~		Queue		7		
~		3		8		
~				9		
~		# of Reads		10	A Comment	
~		Paired End		11	8	
~				12		



Dropin Tab

	2								1			Kit Selection
			How to use:									ForenSeq
				Data' to import da	ta file. Choos	e 'allsequences.txt'	file from R1	or R2 folder				
				ort' or 'Sort and Ex		1-14-14-14-14-14-14-14-14-14-14-14-14-14						
				oly prepares data		'Top20' tab						
						orts top 20 sequence	es for each lo	cus as well a	s the genotype	e for all loci.		
				P. Committee								
												1
									1			(1L)
				This	uorkhook is ra	ited for the followin	or number of	unique cequ	ancer	-		1
			Sample	5000		ited for the followin	ig Hulliber of	unique sequ	-IICC3		_	(1)
			Locus	500								
			Allsequences	50000					-			
			Ansequences	30000								
	12272				Selected Mar							
Locus Active?	Markers	Chromosome		Autosomal	Sex	All						
9,000	AMEL	AUTOSOMAL		Amelogenin	DXS10074	Amelogenin						
•	Amelogenin	AUTOSOMAL		CSF1PO	DXS10103	CSF1PO						
	AMELPGM	AUTOSOMAL	-	D10S1248	DXS10135	D10S1248						
,	CSF1PO	AUTOSOMAL		D12S391	DXS7132	D12S391						
~	D10S1248	AUTOSOMAL		D13S317	DXS7423	D13S317						
0287	D12ATA63	AUTOSOMAL	-	D16S539	DXS8378	D16S539						
-	D12S391	AUTOSOMAL	+	D17S1301	DYF387S1	D17S1301						
	D13S317	AUTOSOMAL	10.5	D18S51	DYS19	D18S51						
Season	D14S1434	AUTOSOMAL		D19S433	DYS385	D19S433						
	D16S539	AUTOSOMAL		D1S1656	DYS3891	D1S1656						
	D17S1301	AUTOSOMAL	+ 30	D20S482	DYS38911	D20S482						
~	D18S51	AUTOSOMAL	- 23	D21S11	DYS390	D21S11						
ν.	D19S433	AUTOSOMAL	13	D22S1045	DYS391	D22S1045						

Dropin Tab

				N .			The state of the s	-11			1/2 2/2 2/2
											Kit Selectio
			How to use:								ForenSeq
						e 'allsequences.txt'	tile from R1 or R2	2 tolder			
				Sort' or 'Sort and Ex	A STATE OF THE PARTY OF THE PAR						
			+	ply prepares data							
			b. Sort and	Export: Both sorts	data and expo	rts top 20 sequence	s for each locus	as well as the gen	otype for all loci.		
										- 1	
									11/1	-	
					1	ted for the followin	ig number of unio	que sequences			1
			Sample	5000							
			Locus	500	1						
			Allsequences	50000							
					Selected Mar	kers					
Locus Active?	Markers	Chromosome		Autosomal	Sex	All					
	AMEL	AUTOSOMAL		1 Amelogenin	DXS10074	Amelogenin					
•	Amelogenin	AUTOSOMAL	-	2 CSF1PO	DXS10103	CSF1PO					
	AMELPGM	AUTOSOMAL	-	3 D10S1248	DXS10135	D10S1248					
,	CSF1PO	AUTOSOMAL		4 D12S391	DXS7132	D12S391					
~	D10S1248	AUTOSOMAL		5 D13S317	DXS7423	D13S317					
	D12ATA63	AUTOSOMAL	-	6 D16S539	DXS8378	D16S539					
~	D12S391	AUTOSOMAL	-	7 D17S1301	DYF387S1	D17S1301					
4	D13S317	AUTOSOMAL		8 D18S51	DYS19	D18S51					
	D14S1434	AUTOSOMAL		9 D19S433	DYS385	D19S433					
~	D16S539	AUTOSOMAL		0 D1S1656	DYS3891	D1S1656					
	D17S1301	AUTOSOMAL	+ 35	1 D20S482	DYS38911	D20S482					
~	D18S51	AUTOSOMAL	1	2 D21S11	DYS390	D21S11					
~	D19S433	AUTOSOMAL	1	3 D22S1045	DYS391	D22S1045					

Process

Include sample	Sample Name		Note: Compile Sample List button allows the user to		Sample List	
~	007A_S6_L001		select the folder containing the STRait Razor Results.	1	007A_S6_L001	E:\Budowle\ISHI 2016\MPS Workshop\007A_S6_L001\R1\allsequences.txt
~	007A_S6_L001	Compile	From there, you can remove unwanted samples	2	007A_S6_L001	E:\Budowle\ISHI 2016\MPS Workshop\007A_S6_L001\R2\allsequences.txt
•	4A_S7_L001	Sample	simply by removing the "check" or "a" from column C	3	4A_S7_L001	E:\Budowle\ISHI 2016\MPS Workshop\4A_S7_L001\R1\allsequences.txt
~	4A_S7_L001	List		4	4A_S7_L001	E:\Budowle\ISHI 2016\MPS Workshop\4A_S7_L001\R2\allsequences.txt
~	005A_S5_L001			5	005A_S5_L001	E:\Budowle\ISHI 2016\MPS Workshop\005A_S5_L001\R1\allsequences.txt
•	005A_S5_L001	Process		6	005A_S5_L001	E:\Budowle\ISHI 2016\MPS Workshop\005A_S5_L001\R2\allsequences.txt
~		Queue		7		
~				8		
~				9		
~		# of Reads		10	A Comment	
v		Paired End		11	8	
~				12	1	



Sequence Pileup

Sample ID Locus	DoC Sequence Data	Nominal Allele
24885 HPRTB	1589 CTAGAACTTATCTTCTGTAAATCTGTCTCTATTTCCATCTCTGTCTCCATCTTTGTCTCTATCTCTATCTCTATCTCTATCT	13 HPRTB [CE 13]-GRCh38-ChrX 134481429-134481588 (TCTA)14
24885 HPRTB	57 CTAGAACTTATCTTCTGTAAATCTGTCTCTATTTCCATCTCTGTCTCCATCTTTGTCTCTATCTCTATCTGTCTATCTCTATC	12 HPRTB [CE 12]-GRCh38-ChrX 134481429-134481588 (TCTA)13
24885 HPRTB	17 CTAGAACTTATCTTCTGTAAATCTGTCTCTATTTCCATCTCTGTCTCCATCTTTGTCTCTATCTCTATCTGTCTATCTCTATC	14 HPRTB [CE 14]-GRCh38-ChrX 134481429-134481588 (TCTA)15
24885 HPRTB	7 CTAGAACTTATCTTCTGTAAATCTGTCTCTATTTCCATCTCTGTCTCCATCTTTGTCCCTATCTCTATCTGTCTATCTCTATC	13 NovelSeq
24885 HPRTB	5 CTAGAACTTATCTTCTGTAAATCTGTCTCTATTTCCATCTCTGTCTCCATCTTTGTCTCTATCTCTATCTGTCTATCTCTATC	13 NovelSeq
24885 HPRTB	4 CTAGAACTTATCTTCTGTAAATCTGTCTCTATTTCCATCTCTGTCTCCATCTTTGTCTCTATCTCTATCTGTCTATCTCTATC	13 NovelSeq
24885 HPRTB	3 CTAGAACTTATCTTCTGTAAATCTGTCTCTATTTCCATCTCTGTCTCCATCTTTGTCTCTATCTCTATCTGTCTATCTCTATC	13 NovelSeq
24885 HPRTB	3 CTAGAACTTATCTTCTGTAAATCTGTCTCTATTTCCATCTCTGTCTCCATCTTTGTCTCTATCTCTATCTGACTATCTCTATC	13 NovelSeq
24885 HPRTB	3 CTAGAACTTATCTTCTGTAAATCTGTCTCTATTTCCATCTCTGTCTCCATCTTTGTCTCTATCTCTATCTGTCTATCTCTATC	13 NovelSeq
24885 HPRTB	3 CTAGAACTTATCTTCTGTAAATCTGTCTCTATTTCCATCTCTGTCTCCATCTTTGTCTCTATCTCTATCTGTCTATCTCTATC	11 HPRTB [CE 11]-GRCh38-ChrX 134481429-134481588 (TCTA)12
24885 HPRTB	2 CTAGAACTCATCTTCTGTAAATCTGTCTCTATTTCCATCTCTGTCTCCATCTTTGTCTCTATCTCTATCTCTATCTCTATC	13 NovelSeq
24885 HPRTB	2 CTAGAACTTATCTTCTGTAAATCTATCTCTATTTCCATCTCTGTCTCCATCTTTGTCTCTATCTCTATCTGTCTATCTCTATC	13 NovelSeq
24885 HPRTB	2 CTAGAACTTATCTTCTGTAAATCTGTCTCTATTTCCATCTCTGTCTCCATCTTTGTCTCCATCTCTATCTGTCTATCTCTATC	13 NovelSeq
24885 HPRTB	2 CTAGAACTTATCTTCTGTAAATCTGTCTCTATTTCCATCTCTGTCTCCATCTTTGTCTCTATCTCTATCTCTATCTCTATC	13 NovelSeq
24885 HPRTB	2 CTAGAACTTATCTTCTGTAAATCTGTCTCTATTTCCATCTCTGTCTCCATCTTTGTCTCTATCTCTATCTGTCTATCTCTATC	13 NovelSeq
24885 HPRTB	2 CTAGAACTTATCTTCTGTAAATCTGTCTCTATTTCCATCTCTGTCTCCATCTTTGTCTCTATCTCTATCTGTCTATCTCTATC	13 NovelSeq
24885 HPRTB	2 CTAGAACTTATCTTCTGTAAATCTGTCTCTATTTCCATCTCTGTCTCCATCTTTGTCTCTATCTCTATCTGTCTATCTCTATC	13 NovelSeq
24885 HPRTB	2 CTAGAACTTATCTTCTGTAAATCTGTCTCTATTTCCATCTCTGTCTCCATCTTTGTCTCTATCTCTATCTGTCTATCTCTATC	13 NovelSeq
24885 HPRTB	2 CTAGAACTTATCTTCTGTAAATCTGTCTCTATTTCCATCTCTGTCTCCATCTTTGTCTCTATCTCTAGCTGTCTATCTCTATC	13 NovelSeq
24885 HPRTB	2 CTAGAACTTATCTTCTGTAAATCTGTCTCTATTTCCATCTCTGTCTCCATCTTTGTCTCTATCTCTATCTCTATCT	13 NovelSeq



Sequence Pileup

Sample ID Locus	DoC Sequence Data	Nominal Allele	
24885 HPRTB	1589 CTAGAACTTATCTTCTGTAAATCTGTCTCTATTTCCATCTCTGTCTCCATCTTTTGTCTCTATCTCTATCTGTCTATCTCTATC	1:	HPRTB [CE 13]-GRCh38-ChrX 134481429-134481588 (TCTA)14
24885 HPRTB	57 CTAGAACTTATCTTCTGTAAATCTGTCTCTATTTCCATCTCTGTCTCCATCTTTTGTCTCTATCTCTATCTGTCTATCTCTATC	12	HPRTB [CE 12]-GRCh38-ChrX 134481429-134481588 (TCTA)13
24885 HPRTB	17 CTAGAACTTATCTTCTGTAAATCTGTCTCTATTTCCATCTCTGTCTCCATCTTTTGTCTCTATCTCTATCTGTCTATCTCTATC	14	HPRTB [CE 14]-GRCh38-ChrX 134481429-134481588 (TCTA)15
24885 HPRTB	7 CTAGAACTTATCTTCTGTAAATCTGTCTCTATTTCCATCTCTGTCTCCATCTTTTGTCCCTATCTCTATCTGTCTATCTCTATC	1:	NovelSeq
24885 HPRTB	5 CTAGAACTTATCTTCTGTAAATCTGTCTCTATTTCCATCTCTGTCTCCATCTTTTGTCTCTATCTCTATCTGTCTATCTCTATC	13	NovelSeq
24885 HPRTB	4 CTAGAACTTATCTTCTGTAAATCTGTCTCTATTTCCATCTCTGTCTCCATCTTTGTCTCTATCTCTATCTGTCTATCTCTATC	13	NovelSeq
24885 HPRTB	3 CTAGAACTTATCTTCTGTAAATCTGTCTCTATTTCCATCTCTGTCTCCATCTTTGTCTCTATCTCTATCTGTCTATCTCTATC	1:	NovelSeq
24885 HPRTB	3 CTAGAACTTATCTTCTGTAAATCTGTCTCTATTTCCATCTCTGTCTCCATCTTTGTCTCTATCTCTATCTGACTATCTCTATC	13	NovelSeq
24885 HPRTB	3 CTAGAACTTATCTTCTGTAAATCTGTCTCTATTTCCATCTCTGTCTCCATCTTTGTCTCTATCTCTATCTGTCTATCTCTATC	1:	NovelSeq
24885 HPRTB	3 CTAGAACTTATCTTCTGTAAATCTGTCTCTATTTCCATCTCTGTCTCCATCTTTTGTCTCTATCTCTATCTGTCTATCTCTATC	11	HPRTB [CE 11]-GRCh38-ChrX 134481429-134481588 (TCTA)12
24885 HPRTB	2 CTAGAACTCATCTTCTGTAAATCTGTCTCTATTTCCATCTCTGTCTCCATCTTTGTCTCTATCTCTATCTGTCTATCTCTATC	1:	NovelSeq
24885 HPRTB	2 CTAGAACTTATCTTCTGTAAATCTATCTCTATTTCCATCTCTGTCTCCATCTTTTGTCTCTATCTCTATCTGTCTATCTCTATC	13	NovelSeq
24885 HPRTB	2 CTAGAACTTATCTTCTGTAAATCTGTCTCTATTTCCATCTCTGTCTCCATCTTTGTCTCCATCTCTATCTGTCTATCTCTATC	1:	NovelSeq
24885 HPRTB	2 CTAGAACTTATCTTCTGTAAATCTGTCTCTATTTCCATCTCTGTCTCCATCTTTTGTCTCTATCTCTATCTGTCTATCTCTATC	13	NovelSeq
24885 HPRTB	2 CTAGAACTTATCTTCTGTAAATCTGTCTCTATTTCCATCTCTGTCTCCATCTTTGTCTCTATCTCTATCTGTCTATCTCTATC	1:	NovelSeq
24885 HPRTB	2 CTAGAACTTATCTTCTGTAAATCTGTCTCTATTTCCATCTCTGTCTCCATCTTTTGTCTCTATCTCTATCTGTCTATCTCTATC	1:	NovelSeq
24885 HPRTB	2 CTAGAACTTATCTTCTGTAAATCTGTCTCTATTTCCATCTCTGTCTCCATCTTTTGTCTCTATCTCTATCTGTCTATCTCTATC	1:	NovelSeq
24885 HPRTB	2 CTAGAACTTATCTTCTGTAAATCTGTCTCTATTTCCATCTCTGTCTCCATCTTTTGTCTCTATCTCTATCTGTCTATCTCTATC	1:	NovelSeq
24885 HPRTB	2 CTAGAACTTATCTTCTGTAAATCTGTCTCTATTTCCATCTCTGTCTCCATCTTTGTCTCTATCTCTAGCTGTCTATCTCTATC	1:	NovelSeq
24885 HPRTB	2 CTAGAACTTATCTTCTGTAAATCTGTCTCTATTTCCATCTCTGTCTCCATCTTTGTCTCTATCTCTATCTCTATCTCTATC	1:	NovelSeq



Genotype Export

Sample ID	Locus	Total Reads	Allele 1	Allele 2	Alt1	Alt2	Alt3	Alt4	Allele 1 DoC	Allele 2 DoC	Alt1 DoC	Alt2 DoC	Alt3 DoC	Alt4 DoC
24885	HPRTB	1846	13		12	14			1589		57	17		

Sample ID	▼ Locus	▼ Unique Seq	▼ Allele ▼	Allele 2	Alt1	Alt2	Alt3	Alt4 ×	Allele 1 DoC ▼	Allele 2 DoC	▼ Alt1 DoC	▼ Alt2 Do	C V Alt3 [DoC - Al	t4 DoC
47A_S9_L001_R2	rs34528025	7_9	10	21	21	21	10	10			32	1	1	1	1
47A_S9_L001_R2	rs2307850	11_12	20		20	20	20	20	50)		1	1	1	1
47A_S9_L001_R2	rs2307700	3_4	10		10	10			53	}		1	1		
47A_S9_L001_R2	rs2307580	2_3	20		20				35	;		1			
47A_S9_L001_R1	rs34528025	7_9	10	21	21	21	10	10	54	1 :	31	1	1	1	1
47A_S9_L001_R1	rs2307850	6_7	20		20	20	20	20	50)		1	1	1	1
47A_S9_L001_R1	rs2307700	4_5	10		10	10	10		54	ı		1	1	1	
47A_S9_L001_R1	rs2307580	2_3	20		20				32	2		1		****	



Allele Export

Sample ID	Locus	Unique Seq	Allele 1	Allele 2	Alt1	Alt2	Alt3	Alt4	Allele 1 SB	Allele 2 SB	Alt1 SB	Alt2 SB	Alt3 SB	Alt4 SB
24885	HPRTB	24_151	CTAGAAG		CTAG	CTAG	SAACTI		0		0	()	

Sample ID	Locus	Total Reads	Allele 1	Allele 2	Alt1	Alt2	Alt3	Alt4	Allele 1 SB	Allele 2 SB	Alt1 SB	Alt2 SB	Alt3 SB	Alt4 SB
47A_S9_L001_R2	rs60901515	60	TGCCTTAT		TGCCTTAT	TGCCTTAT			0.66		0	0		
47A_S9_L001_R2	rs5895446	7	AGGAGAG		AGGAGAC				1		0			
47A_S9_L001_R2	rs4646006	12	TGTAAGTC	TGTAAGT	TGTAAGT	TGTACGTO			0	0	0	0		
47A_S9_L001_R2	rs4187	74	ATGATTAA	ATGATTA	ATGATTA				0.71	0.68	0			
47A_S9_L001_R2	rs3841948	50	AAGTGAT						0.92					
47A_S9_L001_R2	rs3838581	45	GATTACTG	ì	GATTACTO	GATTACTO			0.95		0	0		
47A_S9_L001_R2	rs36062169	65	TTAGGGTT		TTAGGGT	TTAGGGC	TTAGGGT	TTAGGGT	0.88		0	0	0	0
47A_S9_L001_R2	rs36040336	59	CACGGGTT		CACGGGT	CACGGGT	CACGGGT	CACGGGC	0.9		0	0	0	0
47A_S9_L001_R2	rs35769550	98	GCTGCGTT	ACTGCGT	ACTGCGT				0.86	0.88	0			
47A_S9_L001_R2	rs35716687	112	GTCATGCC	GTCATGC	GTCATGC	GTCATGC	GTCATGC	GTCATGC	0.42	0.3	0	0	0	0
47A_S9_L001_R2	rs35605984	58	TATGTCAT		TATGTCAT				0.73		0			
47A_S9_L001_R2	rs34811743	34	TATGTCTC		TATGGCT				0.38		0			
47A_S9_L001_R2	rs34795726	36	ATAATGTA	i	ATAATAT	ATAATGT	ı		0.89		0	0		
47A_S9_L001_R2	rs34541393	74	TACATTTC		TACATTTC	TACATTTC	TACATTTO	TACATTTC	0.79		0	0	0	0
47A_S9_L001_R2	rs34535242	27	AGGGGGT	AGGGGGT	AGGGGGT				0	0	0			



Allele Export (expanded)

Sample ID	Locus	Unique Seq	Allele 1
24885	HPRTR	24 151	CTAGAACTTATCTTCTGTAAATCTGTCTCTATTTCCATCTTGTCTCCCATCTTTGTCTCTCTATCTCTATCTA

Sample ID	Locus	Total Reads	Allele 1
47A_S9_L001_R2	rs60901515	60	TGCCTTATGCAATTTAAGCAACAATAGAAGACAAGTCAGGAACTGAGACTTATCTATTGAAACTCAGGAGTGCTTGGTATCCACAGTGGCAGATAAATTC
47A_S9_L001_R2	rs5895446	7	AGGAGAGATATAGAGTTACTTTGTATCCTGCCACTATCACTGGGGAGATATGTTGGACACAGAGTTCTATCGTGCAAAGTTAAGTGAAAGAGGTTCTAAGGAGATTGTTC
47A_S9_L001_R2	rs4646006	12	TGTAAGTCTAAACAATCAGGCACGTGGGCAGCAATGGAGCTGCAGGTGCACTGTGTGCCATTTACCAGCCTTTGCTGATCTGTTCATTATTTTGCAGGGC
47A_S9_L001_R2	rs4187	74	ATGATTAACAAAAAAAACAAGTAGAAAAATAAGAGAGTGTATTTAAAAAAAA
47A_S9_L001_R2	rs3841948	50	AAGTGATCCAGATTTGGTCTTTTACTGTGAAAATGCTTTTATACAATTTAGTAGAGATGTTATGCAATTGTACTATATCCTTTGCACACTGGAAT
47A_S9_L001_R2	rs3838581	45	GATTACTGGTGTTTACTTTTAATTCCAATAAATTAAAAGTTCTACTGTTTGTT
47A_S9_L001_R2	rs36062169	65	TTAGGGTTTCCTGTCAACTATTCTACTGCCATTTACCACAGGGTCACCACATTCTAATAAGTCCATCCTTCTGAGATATCCTCTTCCTAACATG
47A_S9_L001_R2	rs36040336	59	CACGGGTTAACAGATGCAGTTATTATGCCCATTTAACACGAGGGAAACTGAGGCCCAGAGAGAG
47A_S9_L001_R2	rs35769550	98	GCTGCGTTTCTGTAGAGGAGTAAATGTACTAAGACTATTAAATAACTTACTGACACCTTAACTAAAAACTTTTAGGTGGAAACAAAAGACTGGTTAGAAAAAATG
47A_S9_L001_R2	rs35716687	112	GTCATGCCATCATTAGGGGACTAAATGTGTTAATATCCTGAAAATTATAAGTAATCAATAATTTCTCTTTCGTGATACACCTTGTTTTGAAATATTT
47A_S9_L001_R2	rs35605984	58	TATGTCATAGTAAAAACTTGGAAATAATAAGATGTTGAATAATTGACATTATTAAATTATGCTACATTAGCATAATAAAAATATTAGGTAGTTATTTTTAA
47A_S9_L001_R2	rs34811743	34	TATGTCTCTACATCCCACCCCAACTACAACACTTCGTACCCAGGATGCAACTGAGATCAAAGTAGTTGCTTACTATGGGTTGAACAAAAAGGAGAGGCACAC
47A_S9_L001_R2	rs34795726	36	ATAATGTAGAGTTATTCAAAAAAAAAGGTCTTTCTTTTAGAAATTCTTTTTAAATTATTTGCTACCTATCCATGTTTTCTCCAAATCTATCAGCAGCACAGAGTA
47A_S9_L001_R2	rs34541393	74	TACATTTCTAGATGTGTCAGGAGACTTCAGTTTGGAGAATAAAACTCTACTTCCCTCACATCATTGTTCATACTGTTTTGGTTTTTATTATAA
47A_S9_L001_R2	rs34535242	27	AGGGGGTACTACAGACAGGTTTAAAATGAGCAAACCTAGCTGGTAGGTA



Overview

- How STRait Razor works
- Starting your analysis
- Analysis workbook
- Customization potential
- Try it out



JIT HEALTH SCIENCE CENTER