Repeat Expansion Analysis in GeneMarker<sup>®</sup> software: Streamlined workflow for custom or commercial chemistries of tri- and hexa- nucleotide repeat data, including Huntington's Disease (HTT), Amyotrophic Lateral Sclerosis/Frontotemporal Dementia (ALS, C9ORF72) and Dystrophia Myotonica Protein Kinase (DMPK)

October 2018

Kayla Hendricks, Teresa Snyder-Leiby, Ning Wan SoftGenetics, LLC State College PA

### Introduction

Expansions of simple sequence repeats, mainly but not limited to tri-nucleotide repeats, are responsible for over 40 human diseases.<sup>1</sup> In general, an increasing number of repeats results in more severe phenotype and the number of repeats increase (expand) as the disease gene is inherited.<sup>2</sup>

GeneMarker is a user-friendly tool for rapid and accurate genotyping of repeat expansion data (Figure 1). The new linked Repeat Expansion Application which

- avoids the potentially error prone step of data transfer.
- provides a straight forward user interface to lock in analysis templates that conform to laboratories' standard operating procedures.
- performs the repetitive calculations for converting fragment size to repeat length (Figures 2 and 3).
- print or save final reports with customized header (Figure 4).

# Procedure

1. Import raw data files, make size and allele calls and select Applications – Repeat Expansion Analysis (no need to export sized data to a second analysis software).

**Figure 1:** Link directly to the Repeat Expansion Analysis application from the sized data.

uter: Aug_2018_projectHTT.SGF																	
File View Project Applications Tools Help																	
😂 🕨 🕅 🏋 🇱 Pedigree	×	🔯  [			•   🛛		🖻   🗖	•	Marker:	None	ŀ	•					
E-C Allele Call																	
CAL1_C_10	-07-1	2.11.21.	3708.fsa														
CAL1_C_10		MV	/ Range	1							MW F	Range 2					
CAL1_C_20 Overlay View	00	150	200	250	300	350	400	450	500	550	600	650	700	750	800	850	900
CAL1_C_20E Quantitative Analysis	hma		maina											mmimm			
CALL NETT SNPlex/SNaPshot	1.1																
MSI Analysis																	
Clustering Analysis																	
NA09197_0 Trisomy Analysis	hund																
NA09197_0 Relationship Testing	1		100		200												
NA09197_0 LOH Analysis	12	2	210		200												
MA09197_0. MS-MLPA Analysis	10.0	712111	1 2700	e j													
NA09197_0 Tilling Analysis	10-0	-12-11-2	.1-3700.1														
B NA09197_00 B NA09197_11 Haplotype Analysis	00	150	V Range	1 250	200	250	400	450	500	550	600	cen	700	750	800	950	000
► NA14044 0 ARMS/Comparative Analysis	him	100	200	230	300	330	400	430	300	330	000	0.50	700	730 111111	000	030	500
NA14044 00 Fragile X Analysis																	
NA14044_0 Repeat Expansion Analysis																	
NA14044_020000201010111 20.000	-																
NA14044_02000020100711 20,000 -																	
NA14044_02 20,000   NA14044_050C092018-07-12 20,000   NA14044_0808092018-07-12 10,000																	
NA1404_050092018-07-1; 20,000-   NA1404_050092018-07-1; 10,000-   NA14044_0608092018-07-1; 10,000-   NA14044_0608092018-07-1; 0		Gantraus															

2. Select from a list of analysis templates, or create/modify existing templates	Repeat Expansion Analy	/sis Settings		×
	Template: HTT Hu	ntingtons		:
	Modified	on 2018-10-19 12:27:51	√ s	iave 🗙 Cancel
Figure 2.	Basic information			
Select a template or create one by entering a descriptive	Repeat element:	CAG		
template title (Huntingtons, C9orf72, DMPK)	Description:			
	Genotypes & cutoffs			
	Genotype 🕨	Report category (up to 6 letters)	Cutoff (Repeat# >=)	Report Background
Enter the appropriate Report category Repeat Cutoff values	Normal	NOR	0	
and if desired Highlighting color	Intermediate	NORMu	27 •	
and it desired, frighting color.	Premutation	RedPen	35 -	
	Full Mutation	FulPen	40 .	
		Max Report Repo	at# 200 •	
	Linear calculation facto	N2		
	Correction factor (C	0): 69.715	Ø	Calibrate
Calculate the CO and MO or enter values from lab validation studies	Mobility factor (M0)	2.910	ß	Edit Manually
Calculate the Co and 100 of enter values from ab valuation studies.	Regression info:	Control: CAL1_C_20B112018-07 R^2 = 1.0000	7-12-11-21-3708.fsa.	
	Analyze dye: Blue			
			Ok	Cancel

SoftGenetics LLC 100 Oakwood Ave. Suite 350 State College, PA 16803 USA Phone: 814/237/9340 Fax 814/237/9343 www.softgenetics.com email: info@softgenetics.com



#### 3. Review the results

**Figure 3:** Review results in the application. If background shading was specified in the template, cells in that range will have the designated color. If shading was not specified, there will not be any shading in report table cells. The analyst can select which of the columns to include in the report table.



#### 4. Save/Print Summary and Individual Sample Reports

roject Name: Aug_2018_project# iample #: 80 nalysis Time: 10/22/2018 - 11:34:	нтт 22							
Sample ID		Gen	otype		Peak 1	Peak 2	Peak 3	Peak 4
Sample File	NOR	NORMu	RedPen	FulPen	Repeat1	Repeat1	Repeat2	Repeat2
CAL1_C_10C112018-07-12-11-21-3 708.fsa	×		x	×	17	39	50	75
CAL1_C_10D112018-07-12-11-21-3 708.1sa	×		x	×	17	39	50	75
CAL1_C_20A112018-07-12-11-21-3 708.fsa	x		x	x	17	39	50	75
CAL1_C_20B112018-07-12-11-21-3 708.fsa	×		x	x	17	39	50	75
CAL1_NE112018-07-12-11-21-3708 fsa	×		x	×	17	39	50	75
CAL1_NF112018-07-12-11-21-3708 fsa	×		x	×	17	39	50	75
CAL1_NG112018-07-12-11-21-3708 f8a	×		x	×	17	39	50	75
CAL1_NH112018-07-12-11-21-3708 fea	×		x	×	17	39	50	75
VA09197_001G102018-07-12-11-21 3708.fsa	×			x	18	177		
VA09197_005F102018-07-12-11-21 3708.fsa	×			×	18	176		
IA09197_010E102018-07-12-11-21 3708.fsa	×			×	18	176		
VA09197_020D102018-07-12-11-21 3708.fsa	×			x	18	176		



**Figure 4:** Select the desired reports to print or save for electronic records. Reports are named using the lab specified template name. The summary report provides a quick overview of the project results; listing each sample, x in the cell that corresponds to the peak range(s) for the sample and the calculated repeat number of each fragment.

# Conclusion

The Repeat Expansion Application provides a user-friendly tool to streamline data analysis, customizable templates for different chemistries and reporting flexibility. GeneMarker software is compatible with data files from all major capillary electrophoresis systems (ABI PRISM®, Beckman-Coulter<sup>TM</sup> and MegaBACE<sup>TM</sup>), and Windows® 7 – 10 operating systems.

## References

1. Repeat expansion diseases. Handbook Clin Neurol. 2018;147:105-123. Paulson H.

2. A Brief History of Triplet Repeat Diseases. Helen Budworth and Cynthia T. McMurray Methods Mol Biol. 2013; 1010: 3–17.

### Acknowledgements

Our sincere thanks to the following research scientists for helpful conversations during development of the repeat expansion application: Pamela Snyder (Ohio State University), Ty Lynnes (Indiana University School of Medicine), Michelle Axford (North York General Hospital), and to Robyn Cardwell, Brad Hall, Ninad Pense, Charles Redmond and Sarah Statt (Asuragen Corporation) for helpful discussion and supplying data used in development of the application.

Trademarks are the property of their respective owners. Research Use Only



