Genetic Variants in the *CTGF* Gene Associated with a Family History of Heart Disease

Lorna Smith, Leo Shattuck, Michael Townsend, Jacob Mozier and HE Doherty PhD.

Department of Biological Sciences at Plymouth State University in Plymouth, NH

Introduction

In the United States between 2006 and 2016, deaths due to coronary heart disease have declined by 14.6%. However, problems regarding heart performance may continue throughout the lifespan of surviving individuals. During a heart attack, tissue in the heart becomes damaged, triggering the body to repair the wound. Healthy wound healing involves the gene Connective Tissue Growth Factor (*CTGF*) which is expressed in most major organs in the body, including the heart. *CTGF* also assists in the repair of damaged tissue by signaling cells to lay down structural and support proteins. When overexpressed, *CTGF* can result in the formation of scar tissue. Changes in the base pair sequence of DNA, also known as genetic variants, are hypothesized to impact the function of *CTGF* and therefore the formation of scar tissue. Excess scar tissue can result in loss of electrical conductivity and contractility of heart tissue, leading to chronic heart issues such as increased blood pressure and irregular heartbeat. These heart issues can increase the probability of a repeat heart attack. Our goal is to identify *CTGF* genetic variants in the PSU population. DNA extractions from cheek cells were amplified, sequenced, and compared to the published sequence of the human genome. *CTGF* variants in volunteers were compared with a survey completed at the time of sampling, recording family history of heart disease.

To date we have found 17 genetic variants in the PSU population. Of these 17 variants, 11 have been previously published, and 6 are novel. In the future, increasing the sample size may uncover more *CTGF* variants and reveal the relationship between variants and scarring. If a relationship is found, it may allow for preventative treatments for those most at risk following a heart attack.

Methods

Sampling and Survey

This research was approved by the Plymouth State University IRB. Cheek cells were obtained from PSU volunteers. DNA was extracted using a standard procedure. The *CTGF* gene was amplified via Polymerase Chain Reaction (PCR) and verified using gel electrophoresis. Samples were purified using a QIAquick® PCR purification kit (Qiagen) and sent for sequencing at the Molecular Biology Core Facility at Dartmouth College.

Detecting Variants

Sequences were returned and analyzed for variants using PolyPhred software, where a variant is determined if the base is different from the published sequence from Ensembl.org (GRCh38.p12) and confirmed via visual inspection.

Family Disease Association Scoring

A family history survey was completed by all volunteers at the time of sample collection. To determine a correlation between variant presence and family history, a family history score was calculated using a process similar to Milne et al (2008). Using qualifiers such as obesity, cardiovascular disease, diabetes, high blood pressure, heart attack or stroke before the age of 65 and after 65, family history scores were determined.

- A score of 1 was given for individuals of first-degree (parent, siblings) who have the qualifier.
- A score of 0.5 was given for any second-degree relatives (grandparents) who have the qualifier.

Values from the family history test were compiled to create a cardiovascular disease score for each individual. Kendall's Tau-b and p-values were calculated using a Python program (Van Rossum, 2007). Tau-b values range from -1 to 1, where -1 is a negative correlation, 1 is a positive correlation, and 0 is no correlation. A p-value < 0.021 was considered significant, after a Bonferroni correction for multiple testing.

CTGF Gene Structure

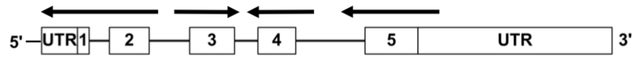

Figure 1: Structure of the *CTGF* gene. Exons are shown as boxes and introns as lines. The locations of the four regions of the gene sequenced in this study and the direction of sequencing are indicated with arrows. UTR – untranslated region.

Detection of Variants

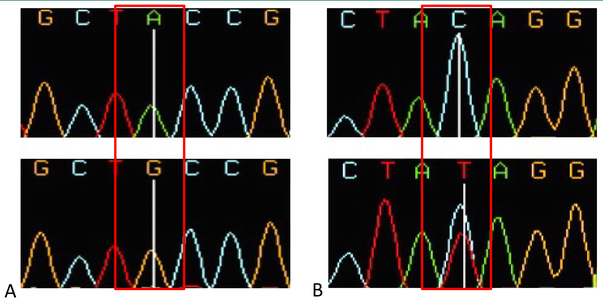

Figure 2: Sanger Sequencing chromatograms showing the difference between A) homozygous and B) heterozygous variants. Homozygous variants appear as a single peak, and heterozygous variants appear as a primary peak, with a secondary peak at the same location.

17 *CTGF* Variants were Identified

Nucleotide Change (- strand)	Type	New/Published	Published Freq.		Sample Freq.	Chromosomes (Sample Size)
			All	CEU		
C156T	5'UTR	New			0.008	3 (365)
C158T	5'UTR	rs112411026	0.001	0	0.005	2 (365)
G485A	C57Y	New			0.018	7 (387)
A534C	Syn	rs6934749	1	1	1	248 (387)
A555C	Syn	rs12206231	1	1	1	248 (387)
G560T	C82F	rs773759696	< 0.001*	N/A	0.036	11 (387)
C562G	H83D	rs7451102	> 0.999	1	1	248 (387)
T596C	V94A	New			0.018	7 (387)
G630A	intron	New			0.012	5 (387)
T774C	intron	rs73779160	0.074	0.015	0.015	5 (333)
C825T	intron (splice)	rs45488997	0.004	0.015	0.015	5 (333)
G1061A	V174M	rs147441296	< 0.001	0	0.003	1 (333)
T1309C	C213R	rs777353866	< 0.001*	N/A	0.007	5 (405)
G1355T	C228F	rs376817540	< 0.001**	N/A	0.145	59 (405)
C2083T	Syn	New			0.016	7 (423)
C2095T	Syn	rs141279872	0.001	0	0.002	1 (423)
T2221C	3'UTR	New			0.016	7 (423)

Table 1: Variants identified in our sample population. Published frequencies were obtained from 1000 Genomes Project data when available where "All" includes all racial groups and "CEU" is the subset of individuals that most closely resemble our own sample population (Utah residents with Northern and Western European ancestry). Other sources of published frequency data are denoted with asterisks.

UTR – Untranslated region, Syn – synonymous.

G1355T is Cardioprotective

A

Variant(s)	Type	Trait	p-value	tau	Variant Frequency
G485A & T596C	C57Y & V94A	Diabetes	< 0.001	0.205	0.018
		MI post 65	0.011	0.139	
T774C	Intron	High BP	< 0.001	0.186	0.015
G1355T	C228F	CVD	0.006	-0.128	0.145

B

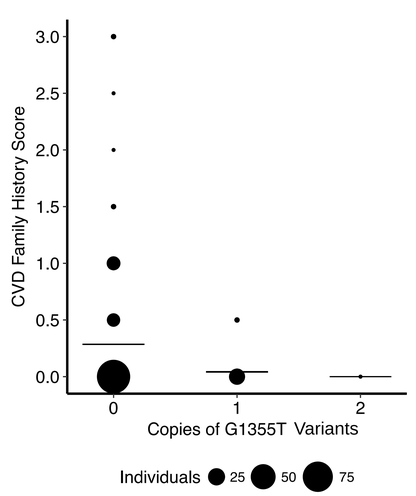

Table 2: Variants correlated with family history of disease. A) Variants with significant Kendall's Tau-b correlation results. Tau-b values range from -1 to 1, where -1 is a negative correlation, 1 is a positive correlation, and 0 is no correlation. B) Bubble plot showing the relationship between number of copies of the G1355T variant and cardiovascular (CVD) disease score.

Conclusions

- Small targeted resequencing efforts in New Hampshire was able to lead to the discovery of several novel genetic variants
- G1355T is common in the population at 16.2%, and may be correlated with a decrease in cardiovascular disease (CVD).

Future Directions

- Continue sequencing the CTGF gene to discover more genetic variants in a broader sample set.
- Determine if genetic variants have an association with family history of a specific heart related disease.

References

1. Ahmed, M. S. *et al.* Connective tissue growth factor--a novel mediator of angiotensin II-stimulated cardiac fibroblast activation in heart failure in rats. *J. Mol. Cell. Cardiol.* **36,** 393–404 (2004).
2. Barrett, J. C., Fry, B., Maller, J. & Daly, M. J. Haploview: analysis and visualization of LD and haplotype maps. *Bioinforma. Oxf. Engl.* **21,** 263–265 (2005).
3. Benjamini, Y., and Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society B.* **57**, 289-300 (1995).
4. Bhangale, T. R., Stephens, M. & Nickerson, D. A. Automating resequencing-based detection of insertion-deletion polymorphisms. *Nat. Genet.* **38,** 1457–1462 (2006).
5. Chen, M. M., Lam, A., Abraham, J. A., Schreiner, G. F. & Joly, A. H. CTGF expression is induced by TGF- beta in cardiac fibroblasts and cardiac myocytes: a potential role in heart fibrosis. *J. Mol. Cell. Cardiol.* **32,** 1805–1819 (2000).
6. Chuva de Sousa Lopes, S. M. *et al.* Connective tissue growth factor expression and Smad signaling during mouse heart development and myocardial infarction. *Dev. Dyn. Off. Publ. Am. Assoc. Anat.* **231,** 542–550 (2004).
7. Dean, R. G. *et al.* Connective Tissue Growth Factor and Cardiac Fibrosis after Myocardial Infarction. *J. Histochem. Cytochem.* **53,** 1245–1256 (2005).
8. Ewing, B., Hillier, L., Wendl, M. C. & Green, P. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res.* **8,** 175–185 (1998).
9. Ewing, B. & Green, P. Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res.* **8,** 186–194 (1998).
10. Fernandez-Rachubinski, F., Eng, B., Murray, W. W., Blajchman, M. A. & Rachubinski, R. A. Incorporation of 7-deaza dGTP during the amplification step in the polymerase chain reaction procedure improves subsequent DNA sequencing. *DNA Seq. J. DNA Seq. Mapp.* **1,** 137–140 (1990).
11. Francis, G. S., McDonald, K., Chu, C. & Cohn, J. N. Pathophysiologic aspects of end-stage heart failure. *Am. J. Cardiol.* **75,** 11A–16A (1995).
12. Frangogiannis, N. G. The inflammatory response in myocardial injury, repair, and remodeling. *Nat. Rev. Cardiol.* **11,** 255–265 (2014).
13. Gerritsen, K. G. *et al.* Plasma CTGF is independently related to an increased risk of cardiovascular events and mortality in patients with atherosclerotic disease: the SMART study. *Growth Factors Chur Switz.* **34,** 149–158 (2016).
14. Gordon, D., Abajian, C. & Green, P. Consed: a graphical tool for sequence finishing. *Genome Res.* **8,** 195–202 (1998).
15. Gordon, D. Viewing and editing assembled sequences using Consed. *Curr. Protoc. Bioinforma.* **Chapter 11,** Unit11.2 (2003).
16. Koitabashi, N. *et al.* Plasma connective tissue growth factor is a novel potential biomarker of cardiac dysfunction in patients with chronic heart failure. *Eur. J. Heart Fail.* **10,** 373–379 (2008).
17. Milne, B. J. *et al.* How should we construct psychiatric family history scores? A comparison of alternative approaches from the Dunedin Family Health History Study. *Psychol. Med.* **38,** 1793–1802 (2008).
18. Musso, M., Bocciardi, R., Parodi, S., Ravazzolo, R. & Ceccherini, I. Betaine, dimethyl sulfoxide, and 7-deaza-dGTP, a powerful mixture for amplification of GC-rich DNA sequences. *J. Mol. Diagn. JMD* **8,** 544–550 (2006).
19. Ohnishi, H. *et al.* Increased expression of connective tissue growth factor in the infarct zone of experimentally induced myocardial infarction in rats. *J. Mol. Cell. Cardiol.* **30,** 2411–2422 (1998).
20. Sun, Y. & Weber, K. T. Infarct scar: a dynamic tissue. *Cardiovasc. Res.* **46,** 250–256 (2000).
21. Sun, Y., Zhang, J. Q., Zhang, J. & Lamparter, S. Cardiac remodeling by fibrous tissue after infarction in rats. *J. Lab. Clin. Med.* **135,** 316–323 (2000).
22. Yang, F. *et al.* Myocardial infarction and cardiac remodelling in mice. *Exp. Physiol.* **87,** 547–555 (2002).