The ClinSeq™ Newsletter

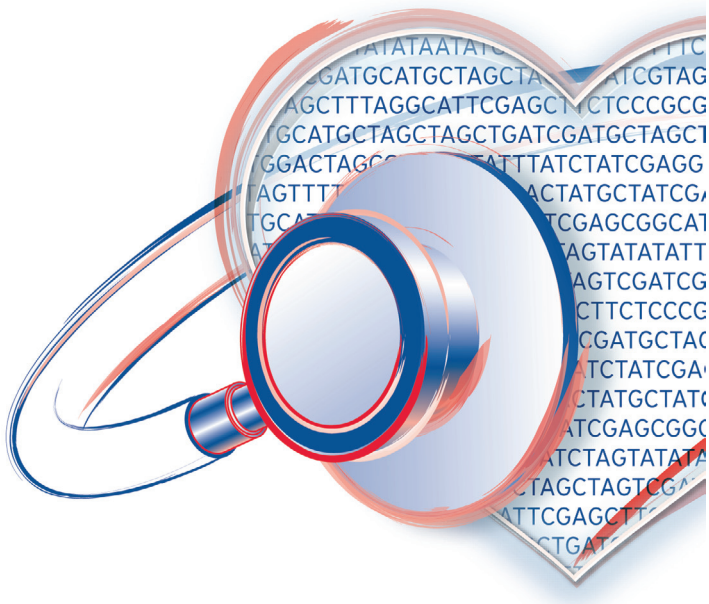
Volume 5, Summer 2012

Inside this issue:

Dear ClinSeq™ Participant,

We are very excited to be entering a new phase of the ClinSeq™ project. Our primary recruitment is coming to a close as we have recruited 958 of our original goal of 1000 participants, and we have begun our "Phase 2 " recruitment, which is specifically directed at enrolling African Americans to increase the diversity of our cohort. We are still recruiting individuals with heart disease of all ethnic groups in order to reach our final goal. Additionally, we are now increasing our efforts to interpret and return results, so more of you will begin hearing from us soon. As always, we are terribly grateful to you for your participation and welcome your input. Thank you for taking the time to look at this update from our work together.

Leslie Biesecker, M.D.
Principal Investigator
Chief, Genetic Disease Research Branch
NHGRI



Contact Information Updates

Are you relocating or changing your phone number? If your phone number(s) or address changes, please let us know. You can call (301) 443-6160 or e-mail clinseq@mail.nih.gov. We need to have your up-to-date contact information so that we can share the latest ClinSeq™ information with you and let you know when genetic results become available for you.

Do you have questions about the study or want to refer a participant? If you need information or have questions about your clinical tests (such as your echocardiogram, EKG, CT scan, or blood work) or the study in general, please contact our research assistant at (301) 443-6160.

Where Are My Results? A Summary of Progress on Returning Results

Many of our participants have been wondering when they can expect to be contacted with the results of their genetic sequencing, and what type of results we have been sharing with participants so far. We are excited to report that we have been steadily analyzing sequence data and are continuing to return results to a growing number of our participants. Thus far, 28 of our 900+ participants have been called back to receive a genetic result. You may ask why the vast majority of people have not gotten information back yet, and the answer lies in the interpretation of the sequence data that we have. As part of the ClinSeq™ project, we have spent several years developing computer programs that can help us analyze the data that we generate through genetic sequencing, and integrate it with information from other databases of genetic information. In addition, it takes a highly skilled staff member to combine your sequence information with data from your personal and family histories and other research studies before we can share a result with you.

Many people have also asked how we decide which results to return first. The answer is that we have begun with the results that have the most serious health implications and those that we feel most confident in—that is, the genetic variations that have been seen previously and have a well-established connection to a disease. To give you a sense of the wide variation in the types of results that we have returned, please see the table on the right. Keep in mind that this is only a fraction of what we intend to return in the coming years, and that we expect that all participants will eventually be contacted with not just one, but multiple results. Also, it's important to know that just because you have not been contacted does not mean that you do not have genetic variation in the genes that we have listed. It may be that your material has not yet been sequenced or reviewed.

Now that our primary recruitment efforts are slowing down, we have started to shift our emphasis towards interpreting and returning results. This should result in a drastic increase in the number of participants who are contacted with the option of receiving results in the next 1-2 years.

Gene with Variation	Number of Participants with Results Returned	Health Implications
LDLR or APOB	12	High cholesterol at a young age that requires medication
BRCA1 or BRCA2	7	Increased risk for cancers including: breast, ovarian, prostate and others
RYR1	2	Malignant hyperthermia, which causes a fast rise in body temperature and severe muscle contractions after a person is given anesthesia
PMP22	3	Numbness or weakness in the limbs
SDHC	1	Mostly benign, but occasionally malignant tumors
KCNH2 and SCN3B	2	Long QT syndrome and Brugada syndrome – serious heart rhythm disorders
PKD1	1	Polycystic kidney disease, which causes cysts in the kidney that can lead to high blood pressure and kidney failure

Analyzing Your Genetic Data

Recent advances in DNA sequencing technology have resulted in dramatically decreased costs for sequencing a human genome, as we are doing for all of our ClinSeq™ participants. While this process cost $75 million dollars per person ten years ago, it now costs $7,500. However, this lower cost presents a surprising challenge to researchers: the huge amount of information that can now be easily generated is challenging to analyze and interpret.

We use several different approaches to analyze and understand this information so that we can generate results to share with you. First, your genetic sequence is processed on computers to identify the positions where your genetic material differs from that of others in the study. We expect that most genetic material (about 99.9%) will be the same across all of our participants. Part of the ClinSeq™ study is to determine how the 0.1% of the genome that is different from one person to the next results in the large variety amongst the health of our participants. This process results in about 100,000 interesting genetic changes for each individual in ClinSeq™. Then, the computers are used to pull together information about each one of those genetic changes from databases throughout the world, and summarize that information in one file.

Once we have the file, the final challenge is to allow ClinSeq™ investigators to easily search and analyze these genetic differences so that they can determine which changes are truly important to your health, and share those with you. To make this calculation, our researchers may take into account as many as 50 different pieces of information for each variant across hundreds of participants. In order to do this more efficiently, we needed a user-friendly program that could handle lots of data and make complex calculations. In response to this need, the ClinSeq™ group developed a graphical desktop software tool called VarSifter that allows researchers and clinicians to search, sort, and filter the genetic differences on their desktop computers in many different ways without having to learn computer programming. This allows them to answer many different research questions

very quickly. For example, when a researcher wants to identify any differences in the LDLR gene (which encodes a protein critical for controlling the levels of LDL, or bad cholesterol) he or she can type in the name of the gene, sort and filter the differences based on whether they are known to cause high cholesterol levels, and then identify the individuals who may have these differences. This has been used to identify and then return this information to study participants.

Tools like VarSifter allow researchers to better understand and interpret the vast amount of genetic information new technologies are able to provide. These tools are primarily designed for research use, but may one day serve as a prototype for software programs, allowing individuals to look at and understand their own genetic information.

Review of Ancillary Studies

Now that ClinSeq™ has enrolled a large number of participants, we have begun to develop many new and exciting research questions that are connected to the original goals of the project, but not central to that mission. We call these "spinoff studies", and their goals range from learning more about whether a specific genetic finding causes disease to surveying participants about gastrointestinal problems to asking about the value of the informed consent process. Here, we profile one such study that we completed recently with the help of our ClinSeq™ participants.

This spinoff project centered around understanding the role of a particular gene, called *GCKR*, in causing diabetes. Several ClinSeq™ participants were found to have variations in their *GCKR* gene invited to come back to the National Institutes of Health Clinical Center (NIHCC) for some additional testing in order to find out if their genetic change was important for their health. It's important to note that we didn't tell the participants the name of the gene or anything about it at the time that they came in for their evaluation. As you probably recall from your consent, we currently don't give out information on genetic changes when it isn't clear whether they cause disease or not, which is why we didn't share this information. In addition, we contacted several ClinSeq™ participants to serve as controls (that is, individuals without any changes in their *GCKR* gene) who were willing to also come in and have the same testing as the individuals with the gene changes so that we could compare the two groups later.

The study found that variations in the *GCKR* gene did have a significant impact on total cholesterol, LDL cholesterol and triglyceride levels. Some of those changes were found to increase cholesterol and triglyceride levels, whereas others decreased those levels. This increase was subtle (small in degree) and was not sufficiently significant to be useful on a clinical level for individual testing or adjustment of treatment. That is, changes in the *GCKR* gene cause a statistically significant effect at the population, but not individual, level. This study did not find a definitive link between variations in the *GCKR* gene and Type 2 Diabetes in the patients clinically, but did show that these changes have a subtle effect on where the protein is located within cells. Participants in the study were not given their specific *GCKR* testing results, but did receive a summary of the results and the journal article that was published about the findings. One participant in the study said, "[Participating in the study] was fascinating for me because there were lots of tests going on and the NIH is a premier place. I'm a science person so [joining the study] was a no brainer for me."

New spinoff studies are starting all the time, and we may be contacting you about one in the near future. It's important for you to know that your participation in those studies is optional and does not affect your participation in ClinSeq™ in any way. We hope that you'll consider participating the next time we contact you about a spinoff study!

ClinSeq™ Scientific Updates

Publications: This year, we have published five new articles that relate to ClinSeq™ and either directly or indirectly use our data.

- Green RC, Berg J, Biesecker LG, Dimmock D, Evans JP, Grody WW, Hegde M, Kahlia S, Korf B, Krantz I, McGuire A, Miller D, Murray M, Nussbaum R, Plon S, Rehm HL, Jacob HJ. Exploring concordance and discordance for return of incidental findings from clinical whole genome sequencing. Genet Med 2012 14:405-410

- Teer JK, Green ED, Mullikin JC, Biesecker LG. VarSifter: Visualizing and analyzing exome-scale sequence variation data on a desktop computer. Bioinformatics 2012 28:599-600.

- Biesecker LG. Opportunities and challenges for the integration of massively parallel genomic sequencing into clinical practice: Lessons from the ClinSeq™ project Genet Med 2012 14:393-398

- Rees MG, Ng D, Ruppert S, Turner C, Beer NL, Swift AJ, Morken MA, Blech I, NISC Comparative Sequencing Program, McCarthy MI, Biesecker LG, Gloyn AL, Collins FS. Phenotypic, Cellular, and Kinetic Correlation of Rare Coding Variants in the Human Glucokinase Regulatory Protein. J Clin Invest 2012 122:205-217

- Sloan JL, Johnston JJ, Manoli I, Chandler RJ, Krause C, Carrillo-Carrasco N, Chandrasekaran SD, O'Brien K,

Hauser NS, Sapp JC, NISC, Barshop BA, Berry S, James PM, Champaigne NL, de Lonlay P, Valayannopoulos V, Geschwind MD, Nyhan WL, Biesecker LG, Venditti CP. Whole exome sequencing identifies mutations in ACSF3 as the cause of combined malonic and methylmalonic aciduria. Nat Genet 2011 43:883-886

Presentations: We've also had the opportunity to give 10 presentations on our study and data in the last year including:

- "Evolution or extinction: Adapting the clinical genetics paradigm." Presidential Plenary Session – The Coming Revolution in Medical Genetics: From Double Helix to Genomics and Back Again. American College of Medical Genetics Annual Meeting. Charlotte, 28 March 2012.
- Cancer gene screening from 572 exomes: Lessons and challenges." American Association of Cancer Research Annual Meeting. Chicago, 31 March 2012.
- "Scaling Genetic Counseling: The Challenges of Extracting Clinically Relevant Data from Whole Genome and Exome Sequencing." National Society of Genetic Counselors Annual Education Conference. San Diego, November 2011.
- "NextGen Sequencing Technologies and Genetic Counseling: Early Experiences, Future Directions." National Society of Genetic Counselors Annual Education Conference. San Diego, November 2011.
- "New opportunities in clinical genomics: Hypothesis-generating clinical research and clinical annotation of genomes." Gordon Research Conference. Newport, July 2011.
- "Transcriptome Profiling of Cardiovascular Disease by Massively Parallel Short-Read DNA Sequencing." NIH Research Festival. Bethesda, 2011
- "Transcriptome profiling in atherosclerosis: a combined analysis of RNA-Seq and microarray data from lymphoblastoid cell lines." NIH Center for Human Inflammation Lecture Series. Bethesda, 2011
- "Transcriptome Profiling of ClinSeq™ Project Participants by Massively Parallel Short-Read DNA Sequencing." Advances in Genome Biology and Technology 201. Marco Island, 2011
- "Identifying pathogenic malignant hyperthermia variants in an unscreened population using whole exome sequencing." The American Society of Human Genetics Annual Genetics Conference. Montreal, October 2011.
- "Detecting and interpreting human genome variation: From populations to the individual." Human Genomics: The next 10 years. La Jolla, February 2011.

Website Updates

In response to suggestions we have received from you, we have begun the process of updating the ClinSeq™ website in order to make it a helpful resource that you can come back to anytime for updates on the project. While the revisions are still a work in progress, we intend to modify the content in the coming months to include all previous ClinSeq™ newsletters, links to our publications and presentations, and any new information about ClinSeq™ in the media. If you have any suggestions or feedback on how we can make the ClinSeq™ website more easily accessible or useful, let us know by calling (301) 443-6160 or e-mailing clinseq@mail.nih.gov.

Featured Associate Investigator: Katie Lewis

1. ***What is your position at NIH?***
   I am the newly hired genetic counselor working on the ClinSeq ™ project. I know that many of you worked closely with Flavia Facio, the study's previous genetic counselor, prior to her departure last year, and I look forward to getting to know you in the coming months and years!
2. ***What motivated you to become involved in the ClinSeq™ Study?***
   I was initially exposed to ClinSeq™ when I was still in my training to become a genetic counselor. I attended the training program that is housed here at the NIH, and clearly remember the day that Flavia came to give a seminar on this exciting, new project called ClinSeq™. At that time, the project was still being conceptualized, but I was fascinated by the complex set of challenges that the study team was dealing with, and the prospect that the project would make a huge impact on the world of genetics. Years later, when I had the opportunity to be directly involved with the project, it was an easy decision for me to make! I love that the project balances research with clinical interactions and care, and am enjoying being a part of a project that is still on the cutting edge of genetics.
3. ***What are your other research interests?***
   One of my main research interests is in the social and behavioral aspects of genetic counseling and testing. This is what I call the "human" piece of our project. I'm particularly interested in learning more from each of you about what motivates you to join a project like ClinSeq™, and how the results we share with you (including what we share and how we share it) affects you and your family.