FDC BEAT

www.fdc.to
Newsletter of the Familial Dilated Cardiomyopathy Project at the University of Miami
Miami, FL, USA Volume XI, Issue I1 July 2010

Gene Discovery Projects in the FDC Research Study: Past, Present and Future!

In this issue we review our past and present genetic findings and inform you about new gene discovery projects planned for the near future. While the goal in the FDC project has always focused on the discovery of genes that cause dilated cardiomyopathy, new technologies have improved our approaches over time. We present here the different methods we have used previously, those we are currently using and those we will be using in the near future. These include linkage analysis, single gene sequencing, exome sequencing, copy number variant analyses, and whole genome sequencing.

Past Genetic Analysis - Linkage

The FDC Study was founded by Ray Hershberger at Oregon Health and Science University in 1993. The project began with the recruitment and enrollment of very large, multi-generational families in which many people (at least 8-10) had a diagnosis of idiopathic dilated cardiomyopathy (IDC), or dilated cardiomyopathy of unknown cause. When multiple family members have IDC, it is called familial dilated cardiomyopathy (FDC).

Linkage analysis. Linkage analysis has been used extensively since the 1980s as an approach to discover genetic cause of disease. Ideally this is done in large families with 8-10 people who have the disease being studied, in our case FDC, and at least as many family members who do not have the disease. Large families improve the statistical power, or the chances of finding the disease-causing mutation in the family. Linkage analysis maps subtle variations in the DNA between family members who have the disease and those who do not. These variations are called markers, which can be thought of as signposts at known locations throughout the genome.

Linkage analysis looks at the genetic markers (signposts) that are shared among family members who have the disease, such as FDC. These are then compared to markers in those family members who do not have FDC. Identifying subtle changes among the genetic markers of DNA is called genotyping. After identifying several hundred markers (like finding signposts) across an individual's entire DNA, this data undergoes linkage analysis. If this is successful, it informs the scientist of the general location of a disease-causing mutation. Linkage analysis does not reveal an exact gene or mutation, but targets an area of the genome for further study. Such an area may contain many genes. Individual genes in that area can then be sequenced to find the specific DNA change responsible for the disease, such as FDC.

In the FDC study in the mid 1990s through the early-mid 2000s, we performed linkage analyses in several large families. Even though new technologies are outpacing this approach, data from the linkage analyses from these families still contribute to our discovery efforts.

Current Genetic Analysis - Sequencing

Single gene sequencing is a technique to read each letter of the gene and determine if the gene is spelled properly. It is like reading a long book, looking for any misspellings in a word or a small string of letters missing or added that changes the meaning. Many changes in the spelling of a gene can be detected with gene sequencing, including those changes in the DNA that can increase risk for DCM in an individual. Although sequencing remains a

Continued on page 2

INSIDE THIS ISSUE:

- Gene Discovery Projects in the FDC Study
  - Exome Sequencing
  - Copy Number Variation
  - Whole Genome Sequencing
- The Role of Genes in Our Bodies
- Staff Updates

key method for most genetic studies, including ours, changes in large pieces of DNA can be missed by sequencing. Other types of major DNA changes, including loss of large sections of DNA containing multiple genes (deletions), additions of DNA (insertions, duplications), or when the DNA is spelled backwards (inversions), can also be missed with gene sequencing.

Resequencing Analysis

Resequencing is a term used when sequencing several genes in DNA from many people. From 2006 to 2010, the FDC Research study has sequenced 15 genes in the DNA specimens from 313 individuals with dilated cardiomyopathy. Mutations in these genes have been previously shown to cause DCM. The official names, or symbols, of the genes we have studied include: *LMNA, PSEN1, PSEN2, MYH7, TNNT2, TCAP, LDB3, SCN5A, CSRP3, MYBPC3, MYH6, TPM1, TNNC1, TNNI3*, and *RBM20*. From the sequences of these 15 genes, we have found mutations in approximately 28% of these 313 individuals. Many of these individuals have family members with dilated cardiomyopathy who are also in our study. Sequencing DNA from these family members has helped us confirm and understand the impact of these genetic findings.

The fact that we have found 28% of genetic cause in our DCM families is a major breakthrough, because in 1993 when the FDC Research Project began, no genetic causes for FDC had been discovered yet! Despite our progress, we still have a great deal of discovery work to do. Our goal is to find the genetic cause in the remaining 72% of these 313 families and in other families in our study. We are pleased with our progress, but we still have an enormous amount of work to do.

The scientific publications from these resequencing studies are available (a list of these publications in the scientific literature can be found at the following weblink: http://www.fdc.to/OurPublications.htm).

One final point: while mutations in more than 30 genes have been found to cause dilated cardiomyopathy, we estimate that these 30 genes explain the genetic cause in only 35-40% of cases of IDC/FDC. We interpret this to suggest that at least 10 to 20 genes that are highly relevant to genetic DCM remain undiscovered.

Next Generation Sequencing (NGS)

Sequencing technology has made dramatic advances in just the past 12-18 months. Whereas previously the largest sequencing instruments could produce sequence data from 96 or 384 fragments of DNA simultaneously, NGS has increased this number to tens of thousands of sequences simultaneously. Thus, it is now possible to efficiently and rapidly sequence (or read) hundreds or thousands of genes at the same time. NGS has revolutionized gene sequencing for both clinical and research purposes. However, sequencing is only the first step, as the genetic code, once read, still needs to be translated and analyzed. This takes enormous computer and software facilities, all of which are keeping pace with NGS.

Using NGS for Exome Sequencing

DNA can be divided into two major categories, coding and noncoding. Coding DNA is the DNA that encodes the make-up of the proteins that function in our body. Coding DNA is found within genes in segments of DNA called exons, which are assembled into the genes that code for our proteins. We have an estimated 18-19,000 genes in our bodies (approximately 180,000 exons), but this coding portion makes up only about 1-2% of our total DNA. A new term, the 'exome' is now used to describe all of the exons together. New technology to examine just these sections, the coding sequence, is called **exome sequencing**. NGS methods are now used for exome sequencing to read the exome, or the entire coding sequence of a person's DNA. In our study we are now using exome sequencing to look for mutations causing DCM.

Unlike linkage analysis of the 1980s-1990s, exome sequencing identifies specific variants within specific genes. This means that potential disease-associated variants discovered with exome sequencing can be examined immediately in other people in the family who also have the disease. This will determine if the genetic variant is seen in all people in the family who have the disease, called segregation with disease, which makes scientists more confident that this genetic variant actually causes DCM in the family. For this reason it is beneficial to have as many people in the family participate in our study as possible, both those who have DCM and those who do not have DCM, especially after a potential DCM variant is discovered.

We have now started exome sequencing in 5 large families. We have exciting preliminary results. We also expect to discover new DCM genes using this approach. We have requested grant funding to study more FDC families with exome sequencing.

Copy Number Variant (CNV) Studies

In addition to exome sequencing, the FDC study is currently using another technique to look for gene mutations that can be missed by sequencing studies. This is called **copy number variant analysis**. Led by Dr. Nadine Norton, the FDC research lab team is looking for large portions of genes that could be deleted, duplicated or placed 'backwards' by mistake into the genome. Any of these could alter the gene, but would be missed with usual sequencing techniques (such as single gene or exome sequencing). With this technique, we also expect to uncover additional gene mutations and find new genes that, when altered, can cause DCM. Our preliminary CNV data are very exciting, and we will have more data available later this year.

The role of genes in our bodies

One can think of our genes as a blueprint for who we are. We have 20,000-25,000 genes in our body that make up our blueprint. Our genes provide a code for making specific proteins, and the proteins are what function in our body. The code in our genes is made up of nucleotides, which you can think of as letters. These letters need to be in a specific pattern to make the right code to function in our body. Any misspelling along the way that alters this code could also alter the protein that is created, which may not work properly in the body. Depending on the gene and the type of variant, or misspelling, involved, this can lead to a particular disease.

The Next Big Thing: Whole Genome Sequencing (WGS)

Finally, you may be wondering why we are not sequencing the whole genome. Sequencing all of the human genome is called **whole genome sequencing**. WGS involves looking at the entire DNA make-up of an individual, which includes the ~2% that makes up the coding portion (the exome), as well as the remaining 98% of human DNA. All of it is called the human genome. We now are focusing our studies on exome sequencing because currently, that is the most cost effective approach for gene discovery. We also note that most of the gene changes (mutations) found thus far in FDC are in genes that affect the make-up of vital heart proteins - and these are encoded by the exome. However, as next generation sequencing improves even more in cost and efficiency, we do imagine that one day we will use whole genome sequencing, where the exome, or coding region, as well as the intronic regions (the noncoding regions) are analyzed for variants that may cause DCM. This technology is too expensive for our research program now, but we may shift to whole genome sequencing within the next few years! These are exciting times!

STAFF UPDATES: Ana Morales, MS, CGC has resigned

Ana Morales, a genetic counselor who worked with our research team since our move to Miami, FL in 2007, has resigned, effective April 20, 2010. We know that many of you have interacted with Ana, either in person or by phone, in the last few years. Anyone who previously had contact with Ana can now contact study genetic counselor **Jill Siegfried**, who will be happy to assist you and answer any questions you might have. **Yves Baptiste,** a research coordinator with the study, is also available to help answer any questions you might have and to enroll additional family members into the study. We greatly appreciated Ana's exceptional assistance, and we wish her the best in her new job!

Jill's Contact Information:
Email jsiegfried@med.miami.edu
Direct: 305-243-8297
Toll-free: 877-800-3430, Option 1

Yves' Contact Information:
Email: ybaptiste2@med.miami.edu
Direct: 305-243-1380
Toll-free: 877-800-3430, Option 3

FDC BULLETIN BOARD

CALL US AT OUR TOLL FREE NUMBER: 877-800-3430

MEDICAL FOLLOW-UP

If anyone in your family is newly diagnosed with heart problems, please let us know. Also, if you or anyone in your family has had heart tests performed, either for follow-up or for the first time, regardless of results, we would be interested in receiving copies. Please contact us and we will send you a medical record release form. If we have sent you medical record release form(s), please send us the completed form(s) as soon as possible.

EMAIL US THROUGH THE "Contact Us" PAGE ON OUR WEBSITE: WWW.FDC.TO

ADD ME TO THE MAILING LIST

If you are not currently a participant in our study, but would like to receive our newsletter, please contact us with your name and address, and we will be pleased to add you to our mailing list.

CONTACT INFORMATION UPDATES

If you have moved and/or have an email address we can contact you at, please call or email us so we can get in touch with you for any follow-up and continue to send you our newsletter.

FDC BEAT Newsletter

FDC BEAT is a publication of the Familial Dilated Cardiomyopathy Project in the Cardiovascular Division at the University of Miami, Miller School of Medicine in Miami, FL. The newsletter is not copyrighted and readers may photocopy its content to share with family members and health care professionals. We welcome your feedback.

The FDC Project Group
Cardiovascular Division
P.O. Box 019132, C-205
Miami, FL, 33101

66495H

ADDRESS SERVICE REQUESTED

TO: