

Sarah Lawrence College

DigitalCommons@SarahLawrence

---

Human Genetics Theses

The Joan H. Marks Graduate Program in  
Human Genetics

---

5-2018

## The Utility of GeneMatcher: a Candidate Gene Database

Halie Holmes

*Sarah Lawrence College*

Laura Fisher

*Sarah Lawrence College*

Follow this and additional works at: [https://digitalcommons.slc.edu/genetics\\_etd](https://digitalcommons.slc.edu/genetics_etd)



Part of the [Genetics Commons](#)

---

### Recommended Citation

Holmes, Halie and Fisher, Laura, "The Utility of GeneMatcher: a Candidate Gene Database" (2018). *Human Genetics Theses*. 48.

[https://digitalcommons.slc.edu/genetics\\_etd/48](https://digitalcommons.slc.edu/genetics_etd/48)

This Thesis - Open Access is brought to you for free and open access by the The Joan H. Marks Graduate Program in Human Genetics at DigitalCommons@SarahLawrence. It has been accepted for inclusion in Human Genetics Theses by an authorized administrator of DigitalCommons@SarahLawrence. For more information, please contact [alester@sarahlawrence.edu](mailto:alester@sarahlawrence.edu).

# The Utility of GeneMatcher: a Candidate Gene Database

Halie Holmes and Laura Fisher

Submitted in partial completion of the Master of Science Degree at Sarah Lawrence College,  
May 7, 2018

## Table of Contents

Key Words: page 2

Abstract: page 2

Introduction: page 3

Research Questions: page 10

Methods: page 10

Results: page 14

Figures: page 15

Discussion: page 18

Study Limitations: page 21

Practice Implications: page 22

Research Recommendations: page 23

Conflict of Interest: page 23

References: page 24

## **Key Words**

GeneMatcher, candidate gene, variant analysis, whole exome sequencing, whole genome sequencing, MatchMaker Exchange, Mendelian disease

## **Abstract:**

As the rate at which whole exome and genome sequencing is used in the clinical and research settings increases, the need for a system to share information obtained from these tests increases as well. GeneMatcher is a website platform that ‘matches’ individuals who are interested in sharing information about the same candidate gene. Using data collected by GeneDx, a clinical diagnostic laboratory, we assessed the outcomes of the matches made between external participants who reached out to GeneDx to discuss findings. Our assessment was made by sorting the outcomes of the discussions about ‘matches’ into three categories: ‘collaboration’, ‘one-off’, and ‘not pursuing’. The categories define the level of information sharing between the requestor and GeneDx and to what extent the discussions moved forward. This methodology was used for both an eight-month retrospective study, along with a four-week prospective study. We found that over half of the matches in both studies resulted in discussions occurring between requestors and the GeneDx ordering clinicians who agreed to communicate about their patients. Approximately one fourth of the matches were categorized as collaborations, meaning a level of data sharing occurred which increased the chances of the candidate gene being classified as a disease causing gene. The results of this study support the benefits of data sharing in regard to furthering discussions and knowledge about the clinical effects of variants in candidate gene.

## **Introduction**

Rare diseases affect 1 in 10 Americans, and many of these diseases have a genetic basis (Institute of Medicine, 2010). Often, individuals with rare diseases face an extensive and stressful challenge to receive a diagnosis; if they find one at all. Whole exome sequencing (WES), a genetic test that sequences approximately 95% of the coding region of the genome, is being increasingly used to find the disease genes at play in many patients with undiagnosed genetic disorders. In the last several years, as whole exome sequencing has been introduced into clinical care, it has been shown to shorten the time to find a diagnosis and reduce costs associated with testing these individuals (Joshi et al., 2016). Although researchers and lab scientists have been able to find the genes and variant(s) responsible for many patients' conditions, the majority of patients who undergo whole exome sequencing fail to receive a diagnosis. The current diagnostic yield is below 35% for most patients' phenotypes (Neveling et al. 2013).

Since WES became available, gene discovery has evolved considerably. Multiple systems have been developed to query searches through stored sequencing data. Although the rate of solving 'Mendelian' disorders has increased with the ability to query genes, a large fraction of patients still remains without a diagnosis because they have variants in candidate disease genes, or genes that are not well understood. Variants in candidate genes may have implications of being pathogenic but have not be previously been associated with human disease. There may be supporting evidence of deleterious effects in model organism data, copy number variant data, tolerance of the gene to sequence variation, tissue and timing specific expression or data regarding the gene function or pathway analysis (Retterer et al., 2016). For

such cases, finding just a single additional unrelated case with a deleterious variant in the same gene and overlapping phenotype may provide enough evidence to attribute a certain phenotype to that gene, enabling diagnoses for those patients (Philippakis et al., 2015). In order to facilitate connections between scientists and clinicians with interest in the same candidate disease genes, sequencing data must be available in a widely-shared format, and improvements must be made in searching for patients or model organisms with variants in specific candidate genes.

GeneMatcher is a freely accessible web-based tool that promotes communication between clinicians, researchers, and individuals who underwent genetic testing, many times WES, that resulted in variants of unknown significance or suspected candidate genes. The Baylor-Hopkins center for Mendelian Genomics created GeneMatcher in 2013 with the goal to connect, or 'match', those who have mutations in the same candidate genes and have similar phenotypes in order to facilitate a better understanding of the function of the genes to determine more appropriate care for those individuals. Matching occurs in a portal where parties can input data through their online account. When two parties are interested in the same candidate gene, they will get "matched", and each will get notified of the other party's interest. It is then the responsibility of those two parties to interact and share the information they are compelled to share (Sobreira et al., 2015).

### *Whole Exome Sequencing*

Whole Exome Sequencing (WES) is a method of genetic testing that evaluates variants in ~20,000 different genes (Cho, M., 2017) in individuals with undiagnosed genetic disorders

(Farwell Hagman et al., 2017). To begin, once variants in genes that are part of the exome are found, the variants are analyzed by the bioinformatics team at the lab. The ~50,000-80,000 variants found (Cho, M., 2017) are prioritized using bioinformatics filtering and manual review to narrow down which variants might be causing the patient's symptoms. Several large studies have demonstrated the overall rates of positive results indicative of diagnosis in patients who have undergone whole exome sequencing are as high as 31% diagnostic yield when trio testing is completed and as low as 22% diagnostic yield when proband-only testing is completed (Retterer et al. 2015; Beaulieu et al. 2014; Lee et al., 2014; Yang et al., 2015).

Variants identified through whole exome sequencing may be reported as positive, negative, or variants of uncertain significance (VUSes). A positive result indicates that a particular variant(s) is causing the patient's symptoms. A negative result indicates that WES did not find a variant(s) that the lab feels is causing the patient's symptoms. A variant of uncertain significance indicates that the particular variant(s) found may be causing the patient's symptoms. Variants of uncertain significance are broken down further into different categories indicating how likely the lab feels they are to be causing the patient's symptoms based on data such as what is known about the gene, segregation in the family, etc (Retterer et al. 2016).

Diagnostic whole exome sequencing is becoming a more commonly ordered test for individuals with undiagnosed genetic disorders because it has the ability to detect dual/multiple diagnoses and conditions with variable expressivity. WES can also test multiple relatives at once to better classify variants found in the proband or symptomatic individual being tested. Finally, in addition to providing a diagnosis for characterized diseases, exome sequencing has the

capacity to provide a diagnosis for newly described diseases and uncover novel candidate genes for disease (Farwell Hagman et al., 2017).

### *Candidate Genes*

Candidate genes are genes whose functions are not well understood. The variants that are predicted to be deleterious in novel candidate genes have not been definitively implicated in human disease, therefore scientists are unsure of the mechanisms or inheritance patterns associated with them. At most, the information known about candidate genes is largely based on only a few human cases (Cho, M., 2017).

Collaboration between clinicians and researchers with interest in the same candidate genes is crucial. Finding just a single additional unrelated case with a deleterious variant in the same candidate gene may provide sufficient evidence to causally implicate the gene, enabling a diagnosis for those patients as well as patients harboring a variant in the same disease gene in the future (Philippakis et al., 2015). In two recent studies of whole exome sequencing cases, ~8% of cases had candidate genes reported (Retterer et al. 2016; Farwell Hagman et al., 2017). In the study by Retterer et al., in which the group reported the diagnostic yield of whole-exome sequencing (WES) in 3,040 consecutive cases, a candidate gene was reported in 24.2% of cases. If all cases with reported candidate genes pooled together, the collaborators would be able to complement each other's strengths and find out more information about candidate genes than any one researcher or healthcare provider alone. The use of whole exome sequencing has resulted in connection of several candidate genes to the appropriate phenotypes. With the increasing amount of information being discovered from whole exome and genome

sequencing, the importance of databases that can connect those interested in the same gene grows exponentially.

### *GeneMatcher*

To enter a gene into GeneMatcher, an individual must generate an account that can be linked with the information they post about the gene of interest. An entry must include at least the name of the gene, which can be entered in by gene symbol, Entrez- or Ensembl-Gene ID. Within the entry there is the option, but not requirement, to include specific variant information by base pair position, phenotypic features of individuals with variant, or diagnosis by OMIM number. GeneMatcher also allows researchers working with model systems to submit orthologous genes. Once the entry is submitted to the database, the submitter is contacted when a match occurs. The submission can be edited and will stay active until entry is deleted. If a match is made, the submitters with the common gene of interest will be put in contact through email. After the initial connection, the responsibility of communication will be put onto the submitters. The entry will remain active even after a match is made in hopes more matches will occur (Sobreira et al., 2015).

GeneMatcher is not a searchable web tool and does not collect or post information that is identifiable. Matches between submitters are private, meaning data pertaining to specific details of successful matches must come directly from submitters. This privacy allows for autonomy of the parties contributing information, but also creates barriers when collecting statistical data of the utility of GeneMatcher (Sobreira et al., 2015).

GeneMatcher is part of the Matchmaker Exchange, a data-sharing tool for information regarding rare diseases. This project helps find diagnoses by connecting multiple databases that collect genotypic and phenotypic data from those with rare disease. Aside from GeneMatcher, other databases inputting data into the Matchmaker Exchange include DECIPHER, MyGene2, AGHA Patient Archive, *matchbox*, Monarch Initiative, and PhenomeCentral. While a database to support other databases may seem redundant, the Matchmaker Exchange allows the organizations to maintain their own practices in data collection from the populations they represent, while preserving the access to a mass of information (Phillipakis et al., 2015). With GeneMatcher's participation in the Matchmaker Exchange program, the number of submissions a single entry has the ability to match with broadens to a larger population, which in turn has increased the likelihood of a match.

Major contributors to the GeneMatcher candidate gene pool include BHCMG and GeneDx. BHCMG, or Baylor-Hopkins Center for Mendelian Genomics, is an NIH-funded center whose goal is to provide DNA sequencing, data analysis, interpretation and infrastructure for dbGaP submission, ELSI (consent) review and publication (BHCMG.org). As of July 2015, BHCMD has entered 180 candidate genes, with 69 successful matches and have published several papers about information from the successful matches (Sobreira et al., 2015). GeneDx is a clinical genetic testing laboratory with a world-renowned whole exome sequencing program and an extensive genetic testing menu specializing in rare and ultra-rare disorders (GeneDx.com). As of February 2018, GeneDx has entered 5,696 submissions and 2,590 genes (M. Cho, personal communication, May 2, 2018). Beside the published data from BHCMD and

internal information from GeneDx, little statistical data has been published from other contributors. The research completed in this thesis project will help fill this void of data.

### **Research Question**

As the rate at which whole exome and genome sequencing is used in the clinical and research settings increases, the need for a system to share information obtained from these tests increases as well. The primary research question is to determine the utility of the GeneMatcher system, a website platform that connects individuals interested in the same candidate genes. By addressing the need for up-to-date statistical information regarding GeneMatcher data, we will be able to present the utility of this system from one laboratory's perspective using GeneDx data. This information will hopefully encourage more clinicians, researchers, and individuals to input their candidate genes into GeneMatcher, allowing the amount of shared information to grow and increasing the likelihood of matching and obtaining a diagnosis.

### **Methods**

When clinicians order WES from GeneDx they are aware that results might include variants in candidate genes. Any time GeneDx WES results in variants in candidate genes they enter the candidate gene and basic data about the type of inheritance (e.g. heterozygous or homozygous) into GeneMatcher. Requestors who enter their own case information into GeneMatcher for the same candidate gene receive a notification that a 'match' was made and they have the option of reaching out to GeneDx to request more information. Each time a request was received to discuss a match with GeneDx it was added to an Microsoft Excel sheet

made by individuals working at GeneDx. Each row of the excel sheet consisted of information about the external and internal cases for a particular request. Columns of the excel sheet included 'gene name', 'date of match', 'requestor name', , 'inheritance pattern' and 'phenotype of patient(s) of requestor', 'name of the internal clinician and/or genetic counselor for which the entry matched', 'phenotype of GeneDx patient(s)', 'GeneDx patient identification number', 'specific variant found by GeneDx', 'assigned GC within GeneDx', any specific notes about the interaction between the requestor and GeneDx, and finally the type of final categorization of the interaction (collaboration, one-off, or not pursuing) for each respective match. Following discussion with the external participant or 'requestor' who had reached out to GeneDx, the requests were then categorized as 'not pursuing', 'one-off', or 'collaboration', based on comparing the cases' similarity in phenotype, inheritance pattern, and previous information collected about the mechanism of the gene.

'Collaboration' is a category in which patient information discussed between the requestor and GeneDx had similarities supporting the conclusion that the candidate gene may have played a role in at least most if not all of the patient symptoms. Additionally, often the requestor and/or GeneDx were aware of additional data such as functional studies or a larger cohort of similar patients which added evidence to the hypothesis the variants in the gene of interest having a pathogenic contribution. GeneDx then reached out to the relevant ordering clinician(s) to notify them of the collaboration opportunity and offered to introduce them to the collaborator to share further patient information. If the clinician(s) agreed to discuss further, they, the requestor, and GeneDx worked together to share data about the genotypic and phenotypic information about their respective patients. Usually this discussion was had not

only to learn more about the potential symptoms caused by variants in the candidate gene but also to possibly publish a paper in order to add the information to the medical/scientific literature. 'One-off' is a category in which the requestor and GeneDx shared general genotypic and phenotypic information about their respective patients through the GeneMatcher system and there was enough overlap to support further information sharing about the candidate gene. GeneDx then reached out to the ordering clinician(s) to notify them of the match and offered to introduce them to the requestor to share further patient information. This category differs from a 'collaboration' in that after communication between the parties no further work was planned to publish a paper or add to the information to the literature about that gene. Also, there was not as strong of genotypic or phenotypic overlap or additional data suggesting strong evidence for the gene's contribution to disease. 'Not pursuing' is a category in which the data obtained from the discussions through GeneMatcher did not overlap in phenotype or inheritance pattern, or did not support the known mechanism of the candidate gene. The requestor was notified that the cases did not have enough similarity in order to conclude that the candidate gene variant(s) was involved in causing the symptoms of both patients. GeneDx did not reach out to the ordering clinician(s) in these cases.

Research for this study was broken down into two tiers. During the first tier, a retrospective study was completed. During the retrospective study, we analyzed the data compiled in the Excel sheet from requests between November 2016 and June 2017. The initial categorizations of each gene were completed by members of the GeneDx team. We then calculated the percentages of matches categorized into 'collaboration', 'one-off', and 'not pursuing'. The amount of publications that came from the matches made in this time frame

were collected from within the GeneDx database. During the second tier, a four-week prospective study was completed using information about the requests that were made on a weekly basis between October 2, 2017 and October 27, 2017. We categorized matches into 'collaboration', 'one-off', or 'not pursuing' categories by comparing the phenotype of the matched requestor's patient to the phenotype of the patients with variants in the same gene in the GeneDx database. We then confirmed the final match type of each case with the GeneDx team after discussions with the requestors and/or clinicians were completed. Finally, we calculated how many matches were categorized as 'collaboration', 'one-off', and 'not pursuing' so that we could use this information to understand how many and what percentage of matches successfully resulted in collaborations and one offs. In addition, we compared this information broken down by weeks.

Quantitative data obtained from the retrospective and prospective studies were analyzed using Microsoft Excel. Separate analysis of the data from the retrospective and prospective studies determined the number of genes in GeneMatcher that matched with GeneDx, which genes were matched with GeneDx, how many cases GeneDx had of the gene with which they matched, the number of requests GeneDx received from clinicians interested in sharing data, and the number of matches that resulted in collaborations between GeneDx and the institution who submitted the gene of interest to GeneMatcher.

Request for Exemption from the Sarah Lawrence College Institutional Review Board (IRB) was submitted on January 3, 2018. An exemption from research was granted by the IRB on January 15, 2018.

## **Results**

During the eight-month retrospective study from November 2016 - June 2017, a total of 544 genes were 'matched' and discussed, of which 146 resulted in collaborations (26.84%), 135 resulted in one-offs (24.82%), and 263 were not pursued (48.35%) (Figure I). In combining 'collaboration' and 'one-off' categories, slightly over half (51.66%) of the matches resulted in GeneDx reaching out to clinicians to continue discussions on the genes of interest, ultimately resulting in data sharing that may impact patient care. A total of four publications resulted from genes that were categorized as collaborations during this time period.

In the four-week period of the prospective study, a total of 98 genes were 'matched' and discussed. Of those matches, 24 went onto become collaborations, 31 were categorized as one-off, and 43 were not pursued. Similarly to the retrospective study, slightly under one half (43.87%) of the matches were not pursued (Figure II). 56.13% of the matches resulted in either 'collaboration' or 'one-off', meaning over half of the matches connected individuals to facilitate discussions that had the ability to affect patient care. A total of nine publications resulted from genes that were categorized as collaborations during this time period.

In the month long prospective study, week one of the study had the least number of matches, while the week four had the most, and weeks two and three had similar amounts of matches to each other (Figure III). We then looked at a snapshot of the retrospective data from October 30, 2016 to November 26, 2016 to provide a comparison from a year prior to the timeframe captured during the prospective data collection of 2017. During the comparable 2016 four-week period from the retrospective dataset, requestors contacted GeneDx about a total of 38 genes through the GeneMatcher system. Of those matches, 12 went on to become

collaborations, 7 were categorized as one-off, and 19 were not pursued (Figure IV). We then compared the amount of cases in each category either 'collaboration', 'one-off', and 'not pursuing' (Figure V).

**Figures**

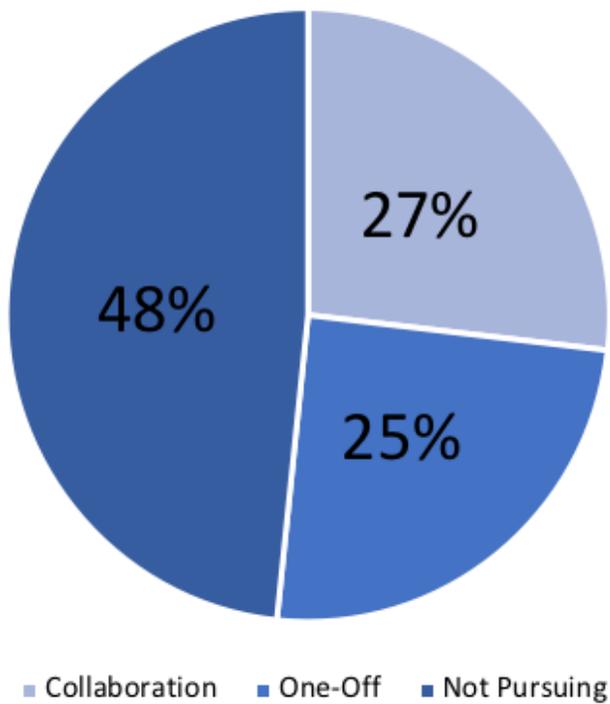


Figure I. Categorization of retrospective matches

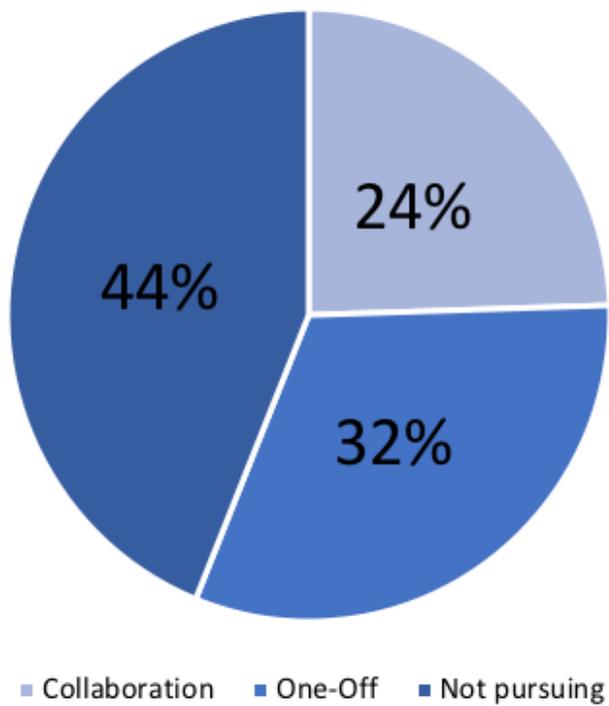


Figure II. Categorization of prospective matches

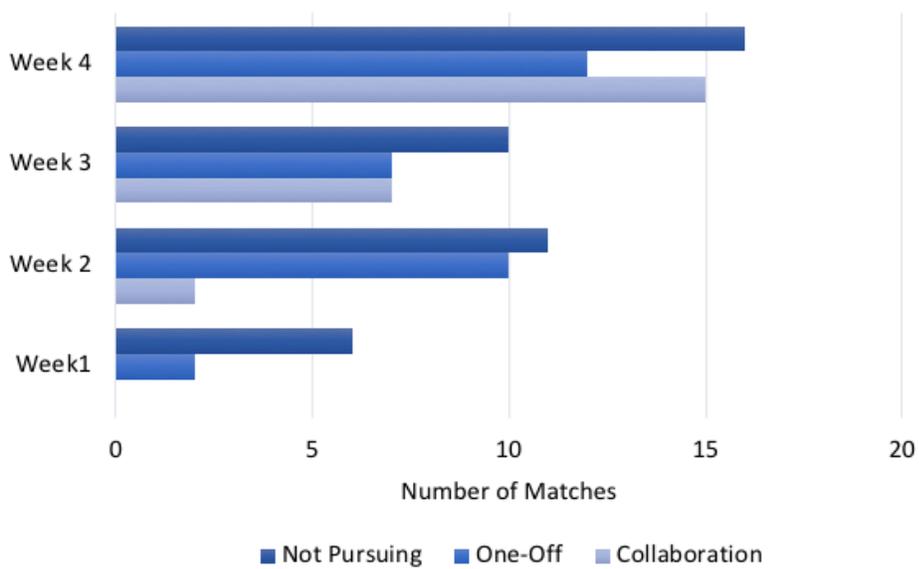


Figure III. Categorization of matches from prospective data separated by week

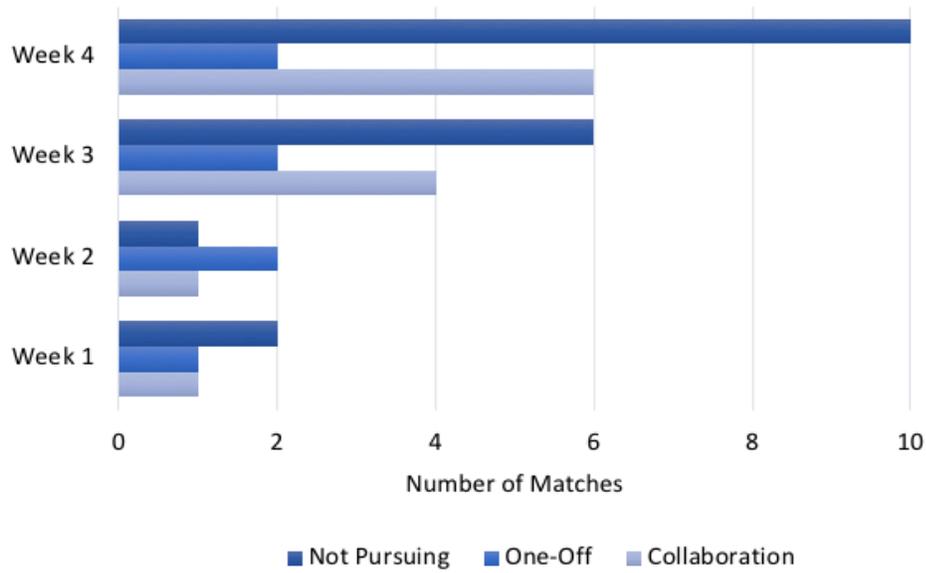


Figure IV. Categorization of matches from retrospective data from October 30, 2016 to November 26, 2016 separated by week

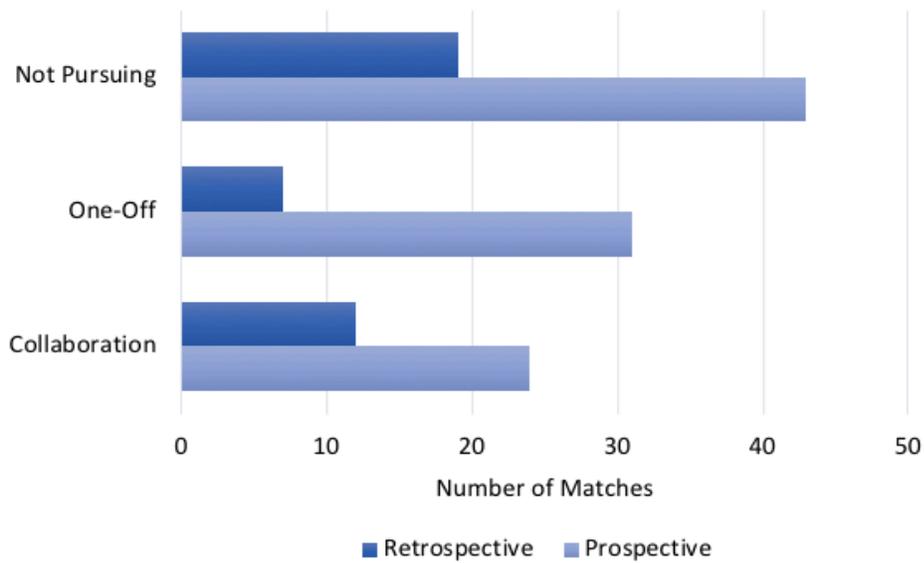


Figure V. Comparison of retrospective data from October 30, 2016 to November 26, 2016 and prospective categorization of matches

## **Discussion**

GeneMatcher is a freely accessible web-based tool created by the Baylor Hopkins Center for Mendelian Genetics in 2013 to connect clinicians, researchers, and individuals interested in the same candidate gene (Sobreira et al., 2015). Variants in candidate genes may have implications such as model organisms data, copy number variant data, tolerance of the gene to sequence variation, tissue and timing specific expression or data regarding the gene function or pathway analysis of being pathogenic but have not be previously been associated with human disease. (Retterer et al., 2016). By connecting the participants interested in the specific candidate gene through GeneMatcher, each are able to share phenotypic and research data to better understand the potential effect of variants in the gene. We know that individuals living without a known genetic diagnosis for themselves or their children feel that having a diagnosis would provide guidance toward medical care and treatments as well as a better view of what the future has in store (Spillman et al., 2017; Lewis et al., 2010; Graungaard and Skov, 2006; Rosenthal et al., 2001). With strong enough evidence, a new syndrome may result from the connections made through the GeneMatcher system, giving those individuals a diagnosis.

The main goal of our study was to quantify the utility of the GeneMatcher system. We examined the occurrence of communication between external clinicians who requesting contact with GeneDx after being 'matched' through GeneMatcher, then analyzed how often the matches resulted data sharing between the parties. By connecting those interested in the same candidate genes, discussions may lead to collaborations to describe new disorders, providing diagnoses, guidance, and support for those living with a previously unknown cause of their disease.

Over the approximately 8-month period from November 2016 - June 2017, over 25% of cases where the requestor reached out to GeneDx after being 'matched' in the GeneMatcher system resulted in 'collaborations', and another approximately 25% resulted in 'one-offs'. Of those reaching out to GeneDx during that time period, there was an approximately 50% chance of being able to discuss the genotype and phenotype further with the GeneDx ordering clinician(s). These connections may have provided enough evidence to attribute certain phenotype to those genes, enabling diagnoses for those patients.

The categorization of matches between the prospective eight-month time period and the prospective one-month time period show similar percentages. In each study, approximately one fourth of the matches were categorized as collaborations. In the retrospective data, one fourth of the results were classified as one-off, while this category in prospective data was closer to one third of the results. There are multiple factors that may have caused the increase in the number of one-offs between the retrospective and prospective studies. One possibility is that more information is known about the particular genes that were 'one-offs' in the prospective study due to the time in which it was completed, which may have allowed the laboratory to compile more information about a larger amount of genes to better inform comparisons in phenotype and inheritance pattern between patients. Another possibility is that during the prospective study, more individuals were involved in categorization, affecting the categorization of the matches and perhaps affecting discernment or human error rates. The not pursuing category was slightly less than one half in each study, 48% in the retrospective and 44% in the prospective. The similarity between the datasets of our study represent the overall consistency of the GeneMatcher system.

When comparing the results of the October 30, 2016 - November 26, 2016 retrospective data to the four weeks of data collected during the prospective study, the first interesting finding is that there are approximately one third of the number of total genes matches in the retrospective data (38 matches) as compared to the prospective data (98 matches). There are multiple factors that may have caused the increase in gene matches. It is likely that more genes were entered into GeneMatcher by both GeneDx as well as other healthcare providers and researchers in 2017 than in 2016 which may be due to possible increased use of WES and awareness of the GeneMatcher system itself. Between the times in which tiers 1 and 2 were completed, the lab's number of overall samples submitted for WES increased from 1422 to 1587, allowing more genes to be entered into the GeneMatcher system and therefore to match with others. As expected, the number of genes categorized as collaboration, one-off, or not pursuing all increased from the four weeks of the retrospective data analyzed in tier 1 to tier 2. In the retrospective data one-month data, one fifth of the results were classified as one-off, while this category in prospective data was closer to one third of the results. The 'not pursuing' category was 50% in the retrospective as compared to 44% in the prospective study. Finally, 32% of the genes matched were collaborations during the October 30, 2016 - November 26, 2016 time period of the retrospective study, as compared to 24% in the prospective study. Although the percentages of collaborations were lower while the one-offs were higher in the month long prospective data, the remaining half of the genes matched fell into the 'not pursuing' category, meaning that approximately half of the matched genes ultimately ended up resulting in data sharing that may impact patient care. This finding is

consistent with the comparison of the overall retrospective data to the prospective data analyzed.

### **Study Limitations**

A limitation to this study is that several individuals were involved in the process of categorizing the inquiries from requestors who genes matched with the GeneDx data into 'collaboration', 'one-off', or 'not pursuing'. Although each requestor inquiry was assessed by only one genetic counselor, the team approach to classifying the inquiries could lead to discrepancies between calls. Another limitation of the study is that if communication between the requestor and GeneDx was cut short before adequate information could be shared to determine if the case was a 'collaboration' or 'one-off', it automatically became categorized as a 'not pursuing'. This could have resulted in inquiries being incorrectly categorized as 'not pursuing', underestimating the distribution of matches that could have been classified as 'one-off' or 'collaboration'. Also matches that were categorized as 'one-off' and 'pursuing' did not always result in further communication, although this was not common. Another limitation is that the retrospective data included data between November 2016 through June 2017, while the prospective data was collected from October 2, 2017 to October 27, 2017. Since the retrospective data and the prospective data were collected during different months of the years, there is no way to account for factors that may affect the amount of data entered into GeneMatcher at those different times. Another limitation is that initial inquiries typically include limited phenotypic data, making it difficult to correctly categorize requests. Also, there are typically small amounts of entries inquiring about the same candidate gene, leading to a

small sample size. Variable expressivity among cases with potentially the same disorder can lead to incorrect conclusions despite best efforts. Another limitation to this study is that the data was collected through a singular lab, GeneDx. This data is not representative of other institutions that enter data into GeneMatcher.

### **Practice Implications**

This study provides statistical data to support the usefulness of the GeneMatcher system. The information obtained in the study gives healthcare providers and researchers a clear understanding of the percentage of genes inputted into the GeneMatcher system that result in an opportunity to share information with other health care providers and researchers. Knowing the usefulness of GeneMatcher encourages individuals to share their data to increase collaborations, leading to the identification of new syndromes or diagnoses. In addition, collaborations between individuals encourages the understanding of the extent that phenotypes may be associated with certain candidate genes. Connecting more individuals with variants in the same rare candidate gene increases the possibility of starting new research or clinical trials. Finally, GeneMatcher helps to create a support network with individuals with variants in the same candidate gene, allowing for the better understanding of the natural history of the disease associated with the candidate gene and for families to connect with each other if desired.

## **Research Recommendations**

Our research recommendations for future research include expanding the size of the dataset. Our prospective data includes data from one-month time period from GeneDx. In the future, we could consider conducting a prospective study over a longer period of time using data from different labs that input data into GeneMatcher and using the same variables or outcomes. We also suggest repeating this study in the future as whole exome sequencing is used more frequently and as the amount of queries in the GeneMatcher system increases. In addition, we suggest comparing the data obtained from GeneMatcher with results from other match making systems when querying the same genes. Finally, we suggest comparing the phenotypes across all individuals with variants in genes that matched in GeneMatcher in order to define which phenotypes are most likely to be ‘matched ‘in GeneMatcher.

## **Conflicts of Interest**

Halie Holmes and Laura Fisher declare they have no conflicts of interest.

## References

- Beaulieu, C. L., Majewski, J., Schwartzentruber, J., Samuels, M. E., Fernandez, B. A., Bernier, F. P., et al. (2014). FORGE Canada Consortium: Outcomes of a 2-Year National Rare-Disease Gene-Discovery Project. *The American Journal of Human Genetics*, *94*(6), 809–817. doi:10.1016/j.ajhg.2014.05.003
- Farwell Hagman, K. D., Shinde, D. N., Mroske, C., Smith, E., Radtke, K., Shahmirzadi, L., et al. (2017). Candidate-gene criteria for clinical reporting: diagnostic exome sequencing identifies altered candidate genes among 8% of patients with undiagnosed diseases. *Genetics in Medicine*, *19*(2), 224–235. doi:10.1038/gim.2016.95
- Graungaard, A. H., & Skov, L. (2007). Why do we need a diagnosis? A qualitative study of parents' experiences, coping and needs, when the newborn child is severely disabled. *Child: Care, Health and Development*, *33*(3), 296–307. <https://doi.org/10.1111/j.1365-2214.2006.00666.x>
- Institute of Medicine. (2010). *Rare Diseases and Orphan Products: Accelerating Research and Development*. <https://doi.org/10.17226/12953>
- Joshi, C., Kolbe, D. L., Mansilla, M. A., Mason, S. O., Smith, R. J. H., Campbell, C. A., ... Campbell, C. A. (2016). Reducing the Cost of the Diagnostic Odyssey in Early Onset Epileptic Encephalopathies, Reducing the Cost of the Diagnostic Odyssey in Early Onset Epileptic Encephalopathies. *BioMed Research International*, *BioMed Research International*, *2016*, *2016*, e6421039. <https://doi.org/10.1155/2016/6421039>, 10.1155/2016/6421039
- Lee, H., Daignan, J. L., Dorrani, N., Strom, S. P., Kantarci, S., Quintero-Rivera, F., ... Nelson, S. F. (2014). Clinical Exome Sequencing for Genetic Identification of Rare Mendelian Disorders. *Jama*, *312*(18), 1880. <https://doi.org/10.1001/jama.2014.14604>
- Neveling, K., Feenstra, I., Gilissen, C., Hoefsloot, L. H., Kamsteeg, E.-J., Mensenkamp, A. R., et al. (2013). A Post-Hoc Comparison of the Utility of Sanger Sequencing and Exome Sequencing for the Diagnosis of Heterogeneous Diseases. *Human Mutation*, *34*(12), 1721–1726. doi:10.1002/humu.22450
- Philippakis, A. A., Azzariti, D. R., Beltran, S., Brookes, A. J., Brownstein, C. A., Brudno, M., ... Rehm, H. L. (2015). The Matchmaker Exchange: A Platform for Rare Disease Gene Discovery. *Human Mutation*, *36*(10), 915–921. <https://doi.org/10.1002/humu.22858>
- Retterer, K., Jussola, J., Cho, M. T., Vitazka, P., Millan, F., Gibellini, F., ... Bale, S. (2016). Clinical application of whole-exome sequencing across clinical indications. *Genetics in Medicine*, *18*(7), 696–704. <https://doi.org/10.1038/gim.2015.148>

- Retterer, K., Scuffins, J., Schmidt, D., Lewis, R., Pineda-Alvarez, D., Stafford, A., et al. (2015). Assessing copy number from exome sequencing and exome array CGH based on CNV spectrum in a large clinical cohort. *Genetics in medicine : official journal of the American College of Medical Genetics*, 17(8), 623–9. doi:10.1038/gim.2014.160
- Rosenthal, E.T., Biesecker, L.G. and Biesecker, B.B. (2001), Parental attitudes toward a diagnosis in children with unidentified multiple congenital anomaly syndromes. *Am. J. Med. Genet.*, 103: 106–114. doi:10.1002/ajmg.1527
- Sobreira, N., Schiettecatte, F., Valle, D., & Hamosh, A. (2015). GeneMatcher: A Matching Tool for Connecting Investigators with an Interest in the Same Gene. *Human Mutation*, 36(10), 928–930. doi:10.1002/humu.22844
- Spillmann, R. C., McConkie-Rosell, A., Pena, L., Jiang, Y.-H., Schoch, K., Walley, N., ... Shashi, V. (2017). A window into living with an undiagnosed disease: illness narratives from the Undiagnosed Diseases Network. *Orphanet Journal of Rare Diseases*, 12(1), 71. <https://doi.org/10.1186/s13023-017-0623-3>
- Yang, Y., Muzny, D. M., Xia, F., Niu, Z., Person, R., Ding, Y., et al. (2014). Molecular Findings Among Patients Referred for Clinical Whole-Exome Sequencing. *JAMA*, 312(18), 1870. doi:10.1001/jama.2014.14601
- Yang, Y., Muzny, D. M., Xia, F., Niu, Z., Person, R., Ding, Y., et al. (2014). Molecular Findings Among Patients Referred for Clinical Whole-Exome Sequencing. *JAMA*, 312(18), 1870. doi:10.1001/jama.2014.14601