



Published in final edited form as:

Hum Mutat. 2022 June ; 43(6): 659–667. doi:10.1002/humu.24373.

Seven years since the launch of the Matchmaker Exchange: the evolution of genomic matchmaking

Kym M. Boycott¹, Danielle R. Azzariti², Ada Hamosh³, Heidi L. Rehm^{2,4}

¹Children's Hospital of Eastern Ontario Research Institute, University of Ottawa, Ottawa, Ontario, Canada

²Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, Massachusetts, USA

³McKusick-Nathans Department of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA.

⁴Center for Genomic Medicine, Massachusetts General Hospital, Boston, Massachusetts, USA

Abstract

The Matchmaker Exchange (MME) was launched in 2015 to provide a robust mechanism to discover novel disease-gene relationships. It operates as a federated network connecting databases holding relevant data using a common application programming interface, where two or more users are looking for a match for the same gene (two-sided matchmaking). Seven years from its launch, it is clear that the MME is making outstanding contributions to understanding the morbid anatomy of the genome. The number of unique genes present across the MME has steadily increased over time; there are currently >13,520 unique genes (~68% of all protein coding genes) connected across the MME's eight genomic matchmaking nodes, GeneMatcher, DECIPHER, PhenomeCentral, MyGene2, *seqr*, Initiative on Rare and Undiagnosed Disease, PatientMatcher, and the RD-Connect Genome-Phenome Analysis Platform. The collective dataset accessible across the MME currently includes more than 120,000 cases from over 12,000 contributors in 98 countries. The discovery of potential new disease-gene relationships is happening daily and international collaborative teams are moving these connections forward to publication, now numbering well over 500. Expansion of data sharing into routine clinical practice by clinicians, genetic counselors and clinical laboratories has ensured access to discovery for even more individuals with undiagnosed rare genetic disease. Tens of thousands of patients and their family members have been directly or indirectly impacted by the discoveries facilitated by two-sided genomic matchmaking. MME supports further connections to the literature (PubCaseFinder) and to human and model organism resources (Monarch Initiative) and scientists (ModelMatcher). Efforts are now underway to explore additional approaches to matchmaking at the gene or variant level where there is only one querier (one-sided matchmaking). Genomic matchmaking has proven its utility over the past 7 years and will continue to facilitate discoveries in the years to come.

Corresponding author: Heidi L. Rehm, Program in Medical and Population Genetics, 105 Broadway, Broad Institute of MIT and Harvard, Cambridge, MA, 02142 USA, hrehm@mgh.harvard.edu.

CONFLICT OF INTERESTS

The authors declare that there are no conflicts of interests.

Keywords

Matchmaker Exchange; rare disease; matchmaking; GA4GH; IRDiRC; novel gene-disease discovery

Introduction

To understand the disease relationships for all genes and variants in the Mendelian genome is nothing short of a grand challenge that represents, and will continue to represent, decades of scientific discovery. The Online Mendelian Inheritance in Man (OMIM) catalog lists 4,224 genes implicated in single gene disorders and traits (omim.org, accessed 3/4/22), yet estimates suggest more than 10,000 genes are likely to be found causal for monogenic disease (Bamshad et al., 2019) and typical diagnostic rates for individuals with suspected genetic disease are only around 25% (100,000 Genomes Project Pilot Investigators et al., 2021). Much more variant-gene-disease discovery remains to be accomplished. As we continue to study increasingly rare Mendelian diseases, the need for responsible data sharing becomes even greater. A key driver of gene-disease discovery for rare disease is the ability to connect two or more parties looking for rare disease cases with the same candidate gene and an overlapping phenotype to expedite the gene-discovery process. This approach has been referred to as two-sided genomic matchmaking and is the predominant use case for the Matchmaker Exchange (MME) (Figure 1) (Philippakis et al., 2015).

The MME consortium was formed in 2013 as a collaborative effort to launch a federated platform to enable two-sided genomic matchmaking through a standardized application programming interface (API) and procedural conventions. While the practice of matchmaking was already occurring within individual databases before the launch of MME, the data was siloed within each node. The MME network set out to connect databases into a federated network, something that had never been done before. The federated MME network launched in 2015 and the work to launch MME was described in detail by Philippakis and colleagues (Philippakis et al., 2015), including formation of a steering committee and developing MME policies and approaches to queries and notifications. The federated MME network API enables exchange of information between the matchmakers through a set of required elements (<https://github.com/ga4gh/mme-apis>). This API is described (Buske, Schiettecatte, et al., 2015), and was developed in collaboration with members of the Global Alliance for Genomics and Health (GA4GH) (Rehm et al., 2021) and the International Rare Diseases Research Consortium (IRDiRC) (Boycott et al., 2017) to ensure interoperability with other genomic services. This early work was described in a special issue of *Human Mutation* in 2015, *The Matchmaker Exchange*, which also included papers describing individual matchmaking services (Buske, Girdea, et al., 2015; Chatzimichali et al., 2015; Gonzalez et al., 2015; Kirkpatrick et al., 2015; Lancaster et al., 2015; Mungall et al., 2015; Sobreira et al., 2015), case reports from clinicians, laboratories and researchers describing novel gene discoveries facilitated by genomic matchmaking (Au et al., 2015; Jurgens et al., 2015; Loucks et al., 2015) as well as other relevant topics (Akle et al., 2015; Brownstein et al., 2015; Krawitz et al., 2015; Lambertson et al., 2015). Seven years later, we now report on

the progress of the MME and advances in the field of genomic matchmaking in this special issue.

Expansion and Impact of the Matchmaker Exchange

When the MME federated network was launched in 2015, it connected three matchmaking services (nodes) (Philippakis et al., 2015): DECIPHER (Database of genomic variation and Phenotype in Humans using Ensembl Resources) (Foreman et al., 2022), GeneMatcher (Hamosh et al., 2022) and PhenomeCentral (Osmond, Hartley, Johnstone, et al., 2022). Since that time, the network (Figure 2) has grown to include five additional genomic matchmaking nodes: MyGene2 (Chong et al., 2016), *seqr* (Pais et al., 2022), Initiative on Rare and Undiagnosed Disease (IRUD) (Adachi et al., 2017), PatientMatcher (Rasi et al., 2022), and the RD-Connect Genome-Phenome Analysis Platform (GPAP) (Laurie et al., 2022) (Table 1). The “Matchmaker Exchange Participants” page on the MME website (www.matchmakerexchange.org/participants.html) provides users a set of tables comparing the types of data stored by each connected MME node, descriptions of matching and scoring output, MME connections, types of users, and summaries of matching and notification protocols. Most of the nodes provide matching services only to users who are storing their genomic data in the database and can flag candidate genes, along with additional information such as phenotype and inheritance pattern, for submission to MME. However, for many interested parties, they are lacking the means and opportunity to contribute this type of comprehensive data so use GeneMatcher (Sobreira et al., 2015) which has the lowest barrier to entry into the MME. As a result, GeneMatcher is the most widely used node of the MME with 11,780 submitters (Hamosh et al., 2022).

The collective dataset accessible across the MME currently includes more than 120,000 cases from over 12,000 contributors in 98 countries. Though it is difficult to track precisely, genomic matchmaking as an approach using one or more nodes of the MME has led to well over 500 publications. The MME provides an accessible approach to matchmaking for individual research labs as well as large-scale research programs. In 2019, Bruel and colleagues (Bruel et al., 2019) reported the outcomes of 2.5 years’ experience using GeneMatcher to share 71 novel candidate genes identified by exome sequencing (ES) and found that the subsequent follow-up of matches supported 39% of genes as causal. Care4Rare Canada has used the MME since its inception and attributes the discovery and publication of 26 novel disease-relationships to connections made via the MME over the last 7 years (unpublished data). Similarly, the Centers for Mendelian Genomics and the Undiagnosed Disease Network both use MME as a critical platform facilitating novel disease-gene discovery (Baxter et al., 2022; Macnamara et al., 2020).

While MME was primarily developed as a platform to service the research community in discovering novel causes of rare disease, it is increasingly being used directly by clinical laboratories that identify candidates through routine clinical genomic sequencing. For example, a *de novo* variant occurring in a highly constrained gene not yet implicated in disease and identified in a trio sent for clinical genomic sequencing would be highly suspicious for novel disease-gene discovery. Entering this gene into MME, along with the phenotype, can often yield matches. This can connect clinical laboratories to investigators

capable of statistical and functional analysis, and ultimately facilitate a diagnosis for a patient. Proof of this can be found in three papers in this special issue report on robust clinical laboratory experiences using MME, including those from Ambry Genetics (Towne et al., 2022), GeneDx (McWalter et al., 2022) and Illumina (Taylor et al., 2022). All three laboratories use GeneMatcher for their matchmaking services; GeneDx submitted entries spanning 3,507 genes, 908 (26%) of which have been validated as causal through evidence built from matchmaking; Ambry Genetics submitted cases spanning 243 unique genes with 111 (45%) now clinically characterized; Illumina has submitted 69 unique genes with 21 (30%) leading to publications or active collaborations with publication planned. Given that some of these genes will be overlapping, it still indicates that, at a minimum, over 900 novel disease genes were identified and validated through MME through just three clinical laboratories using the platform. This number represents more than 20% of the 4,224 genes underlying monogenic disorders and traits cataloged in OMIM (omim.org, accessed 3/4/22). The contributions of clinical laboratories to matchmaking cannot be understated; the 3,507 genes submitted by GeneDx comprised 25% of the 13,941 unique genes in GeneMatcher on 9/28/21 (McWalter et al., 2022).

Seven years from its launch, it is clear that the MME is making outstanding contributions to understanding the morbid anatomy of the genome. The number of unique genes present across the MME has steadily increased over time; there are currently >13,520 genes connected across the MME's eight genomic matchmaking nodes. The discovery of potential new disease-gene relationships is happening daily and tens of thousands of patients and their family members have been directly or indirectly impacted by the discoveries facilitated by two-sided genomic matchmaking.

MME supports connections to knowledge and model organism resources

MME connects to three databases that provide additional utility to the gene discovery scientific community and these databases are considered “Connected Knowledge Sources,” or “Functional Study Node”, specialized MME endpoints that go beyond the initial MME design of patient-patient matchmaking (Figure 2). PubCaseFinder helps users identify any existing case reports for candidate genes by using phenotype-based comparisons (Fujiwara et al., 2022, 2018). To facilitate important downstream translational research using in vitro and in vivo models (Boycott et al., 2020; Wangler et al., 2017), MME connects to two additional databases; Monarch Initiative and ModelMatcher. The Monarch Initiative (Shefchek et al., 2019), supports patient-disease model matches (Mungall et al., 2015), by effectively matching patient phenotype profiles with a potentially existing relevant disease model. ModelMatcher (Harnish et al., 2022) facilitates cross-disciplinary collaborations as part of a global effort to decrease the time to translational and therapeutic research by connecting scientists and other stakeholders who have interest in the same or orthologous gene (Neff, 2021).

Opportunities and Challenges for the Matchmaker Exchange

With greater than 13,520 unique genes connected across the MME, most matchmaking submissions result in at least one match (Figure 3) and each of these matches requires

review that will need additional information to inform the matching parties as to whether or not the match represents a valid novel disease-gene relationship discovery. A matchmaking outcome analysis of 194 genes submitted to MME by the Care4Rare Canada Consortium over a 2-year period highlighted the effectiveness of MME in establishing collaborations for novel candidate genes with a 15% success rate (Osmond, Hartley, Dymont, et al., 2022). The 194 genes entered by Care4Rare Canada resulted in over 1,500 matches returned by MME, with 93% receiving at least one match. Each of these needed to be evaluated, consolidated, and, in the majority of cases, have emails sent to determine the potential significance of the match; this is a laborious process that requires dedicated time and resources. Likewise, *seqr* users have supported manual communication on >6,500 matches (Pais et al., 2022). One of the largest contributing factors to the post-match workload is the lack of phenotypic and inheritance information associated with the majority of cases in GeneMatcher, the largest node of the MME (Osmond, Hartley, Dymont, et al., 2022). To date, only 2.5% of the over 63,000 submissions to GeneMatcher have included phenotype. However, the majority of records in all other primary MME nodes do include this information. The discovery community needs to carefully weigh the pros and cons of a low barrier to entry into MME with the post-match workload.

Moving forward, the MME has the opportunity to improve the efficiency of the matchmaking (Table 2). Inclusion of additional information (phenotype, variant details, zygosity) along with the gene will facilitate the ruling out of matches without the email back-and-forth that is difficult to scale. Osmond and colleagues (2022) showed that ~50% of matches could be ruled out when this type of data was provided by the node at the time of match notification. It is important for users to note that inclusion of high level phenotypic data typically does not require additional consent (Dyke et al., 2017). In addition, the submission of only high quality candidate genes, continuous updating of a group's submissions, removal of cases with ruled out/low priority candidate genes and removal or detailed annotation of solved cases (Table 2) will all promote efficient matchmaking using MME.

Several of the MME nodes have implemented additional features to help users track the large number of matches that currently require follow-up, such as those described for *seqr* (Pais et al., 2022), PhenomeCentral (Osmond, Hartley, Johnstone, et al., 2022), and GeneMatcher (Hamosh et al., 2022) including auto-drafting emails as well as tracking match details, communications and the outcomes of matches within the database. An expanded use case for the MME is the secondary accumulation of additional cases after the initial publication of a novel disease gene relationship. Going forward, opportunities to support connections between investigators made across the MME beyond email notification will be considered, such as the equivalent of a gene 'break-out' room, to help these teams organize with less back-and-forth correspondence.

Evolution of Matchmaking

Matchmaking using MME is based on a two-sided framework where two interested parties are both looking for a match for the same gene (Philippakis et al., 2015). As the past 7 years of experience has demonstrated, this approach has been very successful in advancing

the discovery of novel disease-gene relationships. However, this approach only works when both interested parties have taken the time to flag a highly compelling novel candidate gene of interest. But what of all the datasets where extensive manual review has not occurred due to all sorts of factors? Discovery for these types of datasets needs to happen differently. One-sided matchmaking (Figure 1) can occur when one party is interested in a candidate gene and queries a database hosting genome-wide sequencing data from undiagnosed patients to identify variants in the candidate gene associated with additional information. Zero-sided matchmaking (Figure 1) is the term used to describe the state where there are no candidates identified but instead computational analysis across the cohort is used to identify genes with predicted damaging variants in common across phenotypes. For example, the genebass.org website allows users to query precomputed gene burden analyses across all genes for all phenotypes in the UK Biobank (Karczewski et al., 2021). In another example, the Deciphering Developmental Disorders (DDD) study applies burden testing frameworks to identify genes with significant enrichments of damaging variants, such as genes with more *de novo* loss-of-function variants in the DDD cohort than expected (Kaplanis et al., 2020). Likewise, the new GREGoR consortium (gregorconsortium.org) is amassing rare disease data on the AnVIL platform (Schatz et al., 2022) from both the prior NIH Centers for Mendelian Genomics as well as prospectively collected data to improve power for identifying gene-disease candidates. As more and more data are generated, this type of approach will be critical to ensure we can analyze unsolved datasets at scale.

Several data platforms are approaching one-sided matchmaking by providing information about the existence of a specific variant and its associated information (e.g., phenotype). These databases include MyGene2 (NHGRI/NHLBI University of Washington-Center for Mendelian Genomics (UW-CMG), Seattle, WA), Geno2MP (University of Washington Center for Mendelian Genomics), VariantMatcher (Wohler et al., 2021), and Franklin (Genoox). MyGene2 and Geno2MP are public databases, with sharing driven by families in the case of MyGene2 where anyone can access the displayed variant level data associated with phenotypic terms. VariantMatcher will accept variant-specific queries, search its database of variants, and respond if the variant is present and the associated phenotype if available with dual notification to the querier and data submitter (Wohler et al., 2021). Franklin is an interpretation and connection platform that supports a community of users to facilitate variant interpretation. These four data platforms are working to facilitate a federated connection to one another using Data Connect, a standard for discovery and search from GA4GH (Rodrigues et al., 2022).

At the gene level, several databases, such as DECIPHER (Foreman et al., 2022), RD-Connect GPAP (Laurie et al., 2022), Genomics4RD (Driver et al., 2022), and *seqr* if used in collaboration with the Broad Center for Mendelian Genomics (Pais et al., 2022) are now individually approaching this challenge using internal one-sided matchmaking where an internal user with a candidate gene identified in an undiagnosed patient can query the genomic data housed in the database to see all variants identified in this candidate gene at a certain frequency, or of a certain type, across the dataset along with associated phenotypic and often inheritance data. While these approaches are currently siloed and only available to internal users due to the level of data being shared, efforts are underway to make more of this data available. For example, Geno2MP (University of Washington Center for Mendelian

Genomics) allows searches of the rare variants generated by the majority of the Centers for Mendelian Genomics which are linked to very high-level phenotypic information (Baxter et al., 2022). Genomics4RD (Driver et al., 2022) is piloting a one-sided matchmaking platform for external users using a registered access model to facilitate multi-level filtering for both genetic variation and phenotypic information and ensuring that compound heterozygous variants in a single participant are identifiable. Beacon is a genomic discovery protocol and data access API issued by the GA4GH. Its most recent version (v2) presented in this issue describes its new and enhanced features for complex queries and richer responses (Rambla et al. 2022). Beacon v2 is designed to sit on top of existing solutions, can be integrated into Beacon networks and provides a way forward for the next phase of genomic matchmaking and other data queries.

Concluding Remarks

Over the past seven years, the MME has made outstanding contributions to the discovery of novel disease-gene relationships and is relied on heavily by both the research and clinical rare disease communities. Moving forward, there are opportunities to improve the efficiency of the MME, particularly by encouraging all submitters to share phenotypic and inheritance information with submissions. Although most nodes in MME have from the outset included information about phenotype and inheritance alongside variant submissions, most entries in GeneMatcher, the most widely used node, have not, and therefore we call on the community to quickly move in this direction. Genomic matchmaking approaches continue to evolve and novel approaches to discovery are now underway to ensure that no dataset gets left behind. Connections to literature, model organism resources and scientists, as well as patient-driven matchmaking are all innovative approaches contributing to the ultimate goal of being able to provide diagnostic clarity, biological insight, and social support for the thousands of rare genetic diseases.

ACKNOWLEDGMENTS

The authors thank the members of the MME Consortium for their contributions to this work over the years and their dedication to global data sharing. K.M.B. was supported by a Canadian Institutes of Health Research Foundation grant FDN-154279 and a Tier 1 Canada Research Chair in Rare Disease Precision Health. A.H. was supported by the National Institutes of Health under awards U54HG006542 and RM1HG010860. H.L.R was supported by the National Human Genome Research Institute of the National Institutes of Health under awards U01HG011755, UM1HG008900, and R01HG009141.

WEB RESOURCES

- DECIPHER - <https://decipher.sanger.ac.uk/>
- Franklin - <https://franklin.genoox.com>
- GeneMatcher - <https://www.genematcher.org/>
- Geno2MP - <https://geno2mp.gs.washington.edu/Geno2MP>
- Matchmaker Exchange - <https://www.matchmakerexchange.org/>
- ModelMatcher - <https://www.modelmatcher.net/>
- Monarch Initiative - <https://monarchinitiative.org/>
- MyGene2 - <https://www.mygene2.org/>
- PatientMatcher - <https://www.scilifelab.se/facilities/clinical-genomics-stockholm/>
- PhenomeCentral - <https://www.phenomecentral.org/>

- PubCaseFinder - <https://pubcasefinder.dbcls.jp/>
- RD-Connect GPAP - <https://platform.rd-connect.eu/>
- *seqr* - <https://seqr.broadinstitute.org/matchmaker/matchbox/>
- VariantMatcher - <https://variantmatcher.org/>

REFERENCES

- 100,000 Genomes Project Pilot Investigators, Smedley D, Smith KR, Martin A, Thomas EA, McDonagh EM, Cipriani V, Ellingford JM, Arno G, Tucci A, Vandrovцова J, Chan G, Williams HJ, Ratnaika T, Wei W, Stirrups K, Ibanez K, Moutsianas L, Wielscher M, ... Caulfield M (2021). 100,000 genomes pilot on rare-disease diagnosis in health care - preliminary report. *The New England Journal of Medicine*, 385(20), 1868–1880. 10.1056/NEJMoa2035790 [PubMed: 34758253]
- Adachi T, Kawamura K, Furusawa Y, Nishizaki Y, Imanishi N, Umehara S, Izumi K, & Suematsu M (2017). Japan's initiative on rare and undiagnosed diseases (IRUD): towards an end to the diagnostic odyssey. *European Journal of Human Genetics: EJHG*, 25(9), 1025–1028. 10.1038/ejhg.2017.106 [PubMed: 28794428]
- Akle S, Chun S, Jordan DM, & Cassa CA (2015). Mitigating false-positive associations in rare disease gene discovery. *Human Mutation*, 36(10), 998–1003. 10.1002/humu.22847 [PubMed: 26378430]
- Au PYB, You J, Caluseriu O, Schwartztruber J, Majewski J, Bernier FP, Ferguson M, Care for Rare Canada Consortium, Valle D, Parboosingh JS, Sobreira N, Innes AM, & Kline AD (2015). GeneMatcher aids in the identification of a new malformation syndrome with intellectual disability, unique facial dysmorphisms, and skeletal and connective tissue abnormalities caused by de novo variants in HNRNPK. *Human Mutation*, 36(10), 1009–1014. 10.1002/humu.22837 [PubMed: 26173930]
- Bamshad MJ, Nickerson DA, & Chong JX (2019). Mendelian Gene Discovery: Fast and Furious with No End in Sight. *American Journal of Human Genetics*, 105(3), 448–455. 10.1016/j.ajhg.2019.07.011 [PubMed: 31491408]
- Baxter SM, Posey JE, Lake NJ, Sobreira N, Chong JX, Buyske S, Blue EE, Chadwick LH, Coban-Akdemir ZH, Doheny KF, Davis CP, Lek M, Wellington C, Jhangiani SN, Gerstein M, Gibbs RA, Lifton RP, MacArthur DG, Matise TC, ... O'Donnell-Luria A (2022). Centers for Mendelian Genomics: A decade of facilitating gene discovery. *Genetics in Medicine: Official Journal of the American College of Medical Genetics*. 10.1016/j.gim.2021.12.005
- Boycott KM, Campeau PM, Howley HE, Pavlidis P, Rogic S, Oriel C, Berman JN, Hamilton RM, Hicks GG, Lipshitz HD, Masson J-Y, Shoubridge EA, Junker A, Leroux MR, McMaster CR, Michaud JL, Turvey SE, Dymont D, Innes AM, ... Hieter P (2020). The Canadian Rare Diseases Models and Mechanisms (RDMM) Network: Connecting Understudied Genes to Model Organisms. *American Journal of Human Genetics*, 106(2), 143–152. 10.1016/j.ajhg.2020.01.009 [PubMed: 32032513]
- Boycott KM, Rath A, Chong JX, Hartley T, Alkuraya FS, Baynam G, Brookes AJ, Brudno M, Carracedo A, den Dunnen JT, Dyke SOM, Estivill X, Goldblatt J, Gonthier C, Groft SC, Gut I, Hamosh A, Hieter P, Höhn S, ... Lochmüller H (2017). International Cooperation to Enable the Diagnosis of All Rare Genetic Diseases. *American Journal of Human Genetics*, 100(5), 695–705. 10.1016/j.ajhg.2017.04.003 [PubMed: 28475856]
- Brownstein CA, Holm IA, Ramoni R, Goldstein DB, & Members of the Undiagnosed Diseases Network. (2015). Data sharing in the undiagnosed diseases network. *Human Mutation*, 36(10), 985–988. 10.1002/humu.22840 [PubMed: 26220576]
- Bruel A-L, Vitobello A, Mau-Them FT, Nambot S, Duffourd Y, Quééré V, Kuentz P, Garret P, Thevenon J, Moutton S, Lehalle D, Jean-Marçais N, Orphanomix Physicians' Group, Garde A, Delanne J, Lefebvre M, Lecoquierre F, Trost D, Cho M, ... Thauvin-Robinet C (2019). 2.5 years' experience of GeneMatcher data-sharing: a powerful tool for identifying new genes responsible for rare diseases. *Genetics in Medicine: Official Journal of the American College of Medical Genetics*, 21(7), 1657–1661. 10.1038/s41436-018-0383-z [PubMed: 30563986]
- Buske OJ, Girdea M, Dumitriu S, Gallinger B, Hartley T, Trang H, Misyura A, Friedman T, Beaulieu C, Bone WP, Links AE, Washington NL, Haendel MA, Robinson PN, Boerkoel CF, Adams

- D, Gahl WA, Boycott KM, & Brudno M (2015). PhenomeCentral: a portal for phenotypic and genotypic matchmaking of patients with rare genetic diseases. *Human Mutation*, 36(10), 931–940. 10.1002/humu.22851 [PubMed: 26251998]
- Buske OJ, Schiettecatte F, Hutton B, Dumitriu S, Misyura A, Huang L, Hartley T, Girdea M, Sobreira N, Mungall C, & Brudno M (2015). The Matchmaker Exchange API: automating patient matching through the exchange of structured phenotypic and genotypic profiles. *Human Mutation*, 36(10), 922–927. 10.1002/humu.22850 [PubMed: 26255989]
- Chatzimichali EA, Brent S, Hutton B, Perrett D, Wright CF, Bevan AP, Hurles ME, Firth HV, & Swaminathan GJ (2015). Facilitating collaboration in rare genetic disorders through effective matchmaking in DECIPHER. *Human Mutation*, 36(10), 941–949. 10.1002/humu.22842 [PubMed: 26220709]
- Chong JX, Yu J-H, Lorentzen P, Park KM, Jamal SM, Tabor HK, Rauch A, Saenz MS, Boltshausen E, Patterson KE, Nickerson DA, & Bamshad MJ (2016). Gene discovery for Mendelian conditions via social networking: de novo variants in KDM1A cause developmental delay and distinctive facial features. *Genetics in Medicine: Official Journal of the American College of Medical Genetics*, 18(8), 788–795. 10.1038/gim.2015.161 [PubMed: 26656649]
- Driver HG, Hartley T, Price EM, Turinsky AL, Buske OJ, Osmond M, Ramani AK, Kirby E, Kernohan KD, Couse M, Elrick H, Lu K, Mashouri P, Mohan A, So D, Klamann C, Le HGBH, Herscovich A, Marshall CR, ... Boycott KM (2022). Genomics4RD: An integrated platform to share Canadian deep-phenotype and multi-omic data for international rare disease gene discovery. *Human Mutation*. 10.1002/humu.24354
- Dyke SOM, Knoppers BM, Hamosh A, Firth HV, Hurles M, Brudno M, Boycott KM, Philippakis AA, & Rehm HL (2017). “Matching” consent to purpose: The example of the Matchmaker Exchange. *Human Mutation*, 38(10), 1281–1285. 10.1002/humu.23278 [PubMed: 28699299]
- Foreman J, Brent S, Perrett D, Bevan AP, Hunt SE, Cunningham F, Hurles ME, & Firth HV (2022). DECIPHER: Supporting the interpretation and sharing of rare disease phenotype-linked variant data to advance diagnosis and research. In *Human Mutation*. 10.1002/humu.24340
- Fujiwara T, Shin J-M, & Yamaguchi A (2022). Advances in the development of PubCaseFinder, including the new application programming interface and matching algorithm. *Human Mutation*. 10.1002/humu.24341
- Fujiwara T, Yamamoto Y, Kim J-D, Buske O, & Takagi T (2018). PubCaseFinder: A Case-Report-Based, Phenotype-Driven Differential-Diagnosis System for Rare Diseases. *American Journal of Human Genetics*, 103(3), 389–399. 10.1016/j.ajhg.2018.08.003 [PubMed: 30173820]
- Genoox. (n.d.). Franklin. Franklin by Genoox. Retrieved February 28, 2022, from <https://franklin.genoox.com>
- Gonzalez M, Falk MJ, Gai X, Postrel R, Schüle R, & Zuchner S (2015). Innovative genomic collaboration using the GENESIS (GEM.app) platform. *Human Mutation*, 36(10), 950–956. 10.1002/humu.22836 [PubMed: 26173844]
- Hamosh A, Wohler E, Martin R, Griffith S, da S Rodrigues E, Antonescu C, Doheny KF, Valle D, & Sobreira N (2022). The impact of GeneMatcher on international data sharing and collaboration. *Human Mutation*. 10.1002/humu.24350
- Harnish JM, Li L, Rogic S, Poirier-Morency G, Kim S-Y, Network UD, Boycott KM, Wangler MF, Bellen HJ, Hieter P, Pavlidis P, Liu Z, & Yamamoto S (2022). ModelMatcher: A scientist-centric online platform to facilitate collaborations between stakeholders of rare and undiagnosed disease research. *Human Mutation*. 10.1002/humu.24364
- Jurgens J, Sobreira N, Modaff P, Reiser CA, Seo SH, Seong M-W, Park SS, Kim OH, Cho T-J, & Pauli RM (2015). Novel COL2A1 variant (c.619G>A, p.Gly207Arg) manifesting as a phenotype similar to progressive pseudorheumatoid dysplasia and spondyloepiphyseal dysplasia, Stanescu type. *Human Mutation*, 36(10), 1004–1008. 10.1002/humu.22839 [PubMed: 26183434]
- Kaplanis Joanna, Samochoa Kaitlin E., Wiel Laurens, Zhang Zhancheng, Arvai Kevin J., Eberhardt Ruth Y., Gallone Giuseppe, Lelieveld Stefan H., Martin Hilary C., McRae Jeremy F., Short Patrick J., Torene Rebecca I., de Boer Elke, Danecsek Petr, Gardner Eugene J., Huang Ni, Lord Jenny, Martincorena Iñigo, Pfundt Rolph, Reijnders Margot R. F., Yeung Alison, Yntema Helger G., DDD Study consortium authors, Vissers Lisenka E. L. M., Juusola Jane, Wright Caroline F., Brunner Han G., Firth Helen V., FitzPatrick David R., Barrett Jeffrey C., Hurles Matthew E.,

- Gilissen Christian, Retterer Kyle. Evidence for 28 genetic disorders discovered by combining healthcare and research data. *Nature*. 2020 Oct 1; 586(7831): 757–762. [PubMed: 33057194]
- Karczewski K, Solomonson M, Chao KR, Goodrich JK, Tiao G, Lu W, Riley-Gillis B, Tsai E, Kim HI, Zheng X, Rahimov F, Esmaeeli S, Grundstad AJ, Reppell M, Waring J, Jacob H, Sexton D, Bronson PG, Chen X, ... Neale BM (2021). Systematic single-variant and gene-based association testing of 3,700 phenotypes in 281,850 UK Biobank exomes. In *bioRxiv. medRxiv*. 10.1101/2021.06.19.21259117
- Kirkpatrick BE, Riggs ER, Azzariti DR, Miller VR, Ledbetter DH, Miller DT, Rehm H, Martin CL, Faucett WA, & ClinGen Resource. (2015). GenomeConnect: matchmaking between patients, clinical laboratories, and researchers to improve genomic knowledge. *Human Mutation*, 36(10), 974–978. 10.1002/humu.22838 [PubMed: 26178529]
- Krawitz P, Buske O, Zhu N, Brudno M, & Robinson PN (2015). The genomic birthday paradox: how much is enough? *Human Mutation*, 36(10), 989–997. 10.1002/humu.22848 [PubMed: 26239817]
- Lambertson KF, Damiani SA, Might M, Shelton R, & Terry SF (2015). Participant-driven matchmaking in the genomic era. *Human Mutation*, 36(10), 965–973. 10.1002/humu.22852 [PubMed: 26252162]
- Lancaster O, Beck T, Atlan D, Swertz M, Thangavelu D, Veal C, Dagleish R, & Brookes AJ (2015). Cafe Variome: general-purpose software for making genotype-phenotype data discoverable in restricted or open access contexts. *Human Mutation*, 36(10), 957–964. 10.1002/humu.22841 [PubMed: 26224250]
- Laurie S, Piscia D, Matalonga L, Corvo A, Garcia C, Fernandez-Callejo M, Hernandez C, Luengo C, Ntalis AP, Protassio J, Martinez I, Pico D, Thompson R, Tonda R, Bayes M, Bullich G, Camps J, Paramonov I, Trotta J-R, ... Beltran S (2022). The RD-Connect Genome-Phenome Analysis Platform: Accelerating diagnosis, research, and gene discovery for rare diseases. *Human Mutation*. 10.1002/humu.24353
- Loucks CM, Parboosingh JS, Shaheen R, Bernier FP, McLeod DR, Seidahmed MZ, Puffenberger EG, Ober C, Hegele RA, Boycott KM, Alkuraya FS, & Innes AM (2015). Matching two independent cohorts validates DPH1 as a gene responsible for autosomal recessive intellectual disability with short stature, craniofacial, and ectodermal anomalies. *Human Mutation*, 36(10), 1015–1019. 10.1002/humu.22843 [PubMed: 26220823]
- Macnamara EF, D'Souza P, Undiagnosed Diseases Network, & Tiftt CJ (2020). The undiagnosed diseases program: Approach to diagnosis. *Translational Science of Rare Diseases*, 4(3–4), 179–188. 10.3233/TRD-190045 [PubMed: 32477883]
- McWalter K, Torti E, Morrow M, Juusola J, & Retterer K (2022). Discovery of over 200 new and expanded genetic conditions using GeneMatcher. *Human Mutation*. 10.1002/humu.24351
- Mungall CJ, Washington NL, Nguyen-Xuan J, Condit C, Smedley D, Köhler S, Groza T, Shefchek K, Hochheiser H, Robinson PN, Lewis SE, & Haendel MA (2015). Use of model organism and disease databases to support matchmaking for human disease gene discovery. *Human Mutation*, 36(10), 979–984. 10.1002/humu.22857 [PubMed: 26269093]
- Neff EP (2021). Model matchmaking. *Lab Animal*, 50(2), 39–42. 10.1038/s41684-020-00706-7 [PubMed: 33473233]
- NHGRI/NHLBI University of Washington-Center for Mendelian Genomics (UW-CMG), Seattle, WA. (n.d.). MyGene2. MyGene2. Retrieved February 28, 2022, from <https://mygene2.org/MyGene2/>
- Osmond M, Hartley T, Dyment DA, Kernohan KD, Brudno M, Buske OJ, Innes AM, Boycott KM, & Care4Rare Canada Consortium. (2022). Outcome of over 1500 matches through the Matchmaker Exchange for rare disease gene discovery: The 2-year experience of Care4Rare Canada. *Genetics in Medicine: Official Journal of the American College of Medical Genetics*, 24(1), 100–108. 10.1016/j.gim.2021.08.014 [PubMed: 34906465]
- Osmond M, Hartley T, Johnstone B, Andjic S, Girdea M, Gillespie M, Buske O, Dumitriu S, Koltunova V, Ramani A, Boycott KM, & Brudno M (2022). PhenomeCentral: 7 years of rare disease matchmaking. *Human Mutation*. 10.1002/humu.24348
- Pais LS, Snow H, Weisburd B, Zhang S, Baxter SM, DiTroia S, O'Heir E, England E, Chao KR, Lemire G, Osei-Owusu I, VanNoy GE, Wilson M, Nguyen K, Arachchi H, Phu W, Solomonson M, Mano S, O'Leary M, ... O'Donnell-Luria A (2022). *seqr*: a web-based analysis and collaboration tool for rare disease genomics. In *Human Mutation*. 10.1002/humu.24366

- Philippakis AA, Azzariti DR, Beltran S, Brookes AJ, Brownstein CA, Brudno M, Brunner HG, Buske OJ, Carey K, Doll C, Dumitriu S, Dyke SOM, den Dunnen JT, Firth HV, Gibbs RA, Girdea M, Gonzalez M, Haendel MA, Hamosh A, ... Rehm HL (2015). The Matchmaker Exchange: a platform for rare disease gene discovery. *Human Mutation*, 36(10), 915–921. 10.1002/humu.22858 [PubMed: 26295439]
- Rasi C, Nilsson D, Magnusson M, Lesko N, Lagerstedt-Robinson K, Wedell A, Lindstrand A, Wirta V, & Stranneheim H (2022). PatientMatcher: a customizable Python-based open-source tool for matching undiagnosed rare disease patients via the Matchmaker Exchange network. *Human Mutation*. 10.1002/humu.24358
- Rehm HL, Page AJH, Smith L, Adams JB, Alterovitz G, Babb LJ, Barkley MP, Baudis M, Beauvais MJS, Beck T, Beckmann JS, Beltran S, Bernick D, Bernier A, Bonfield JK, Boughtwood TF, Bourque G, Bowers SR, Brookes AJ, ... Birney E (2021). GA4GH: International policies and standards for data sharing across genomic research and healthcare. *Cell Genomics*, 1(2). 10.1016/j.xgen.2021.100029
- Rodrigues E. da S., Griffith S, Martin R, Antonescu C, Posey JE, Coban-Akdemir Z, Jhangiani SN, Doheny KF, Lupski JR, Valle D, Bamshad MJ, Hamosh A, Sheffer A, Chong JX, Einhorn Y, Cupak M, & Sobreira N (2022). Variant-level matching for diagnosis and discovery: challenges and opportunities. *Human Mutation*. 10.1002/humu.24359
- Schatz MC, Philippakis AA, Afgan E, Banks E, Carey VJ, Carroll RJ, Culotti A, Ellrott K, Goecks J, Grossman RL, Hall IM, Hansen KD, Lawson J, Leek JT, Luria AO, Mosher S, Morgan M, Nekrutenko A, O'Connor BD, ... Wuichet K (2022). Inverting the model of genomics data sharing with the NHGRI Genomic Data Science Analysis, Visualization, and Informatics Lab-space. *Cell Genomics*, 2(1). 10.1016/j.xgen.2021.100085
- Shefchek KA, Harris NL, Gargano M, Matentzoglou N, Unni D, Brush M, Keith D, Conlin T, Vasilevsky N, Zhang XA, Balhoff JP, Babb L, Bello SM, Blau H, Bradford Y, Carbon S, Carmody L, Chan LE, Cipriani V, ... Osumi-Sutherland D (2019). The Monarch Initiative in 2019: an integrative data and analytic platform connecting phenotypes to genotypes across species. *Nucleic Acids Research*. 10.1093/nar/gkz997
- 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. 10.1002/humu.22844 [PubMed: 26220891]
- Taylor JP, Malhotra A, Burns NJ, Clause AR, Brown CM, Burns BT, Chandrasekhar A, Schlachetzki Z, Bennett M, Thorpe E, Taft RJ, Perry DL, & Coffey AJ (2022). A clinical laboratory's experience using GeneMatcher - building stronger gene-disease relationships. *Human Mutation*. 10.1002/humu.24356
- Towne MC, Rossi M, Wayburn B, Huang JM, Radtke K, Alcaraz W, Farwell Hagman KD, & Shinde DN (2022). Diagnostic testing laboratories are valuable partners for disease gene discovery: 5-year experience with GeneMatcher. *Human Mutation*. 10.1002/humu.24342
- University of Washington Center for Mendelian Genomics. (n.d.). Geno2MP. Genotype to Mendelian Phenotype (Geno2MP v2.6) Browser. Retrieved February 28, 2022, from <https://geno2mp.gs.washington.edu/Geno2MP>
- Wangler MF, Yamamoto S, Chao H-T, Posey JE, Westerfield M, Postlethwait J, Members of the Undiagnosed Diseases Network (UDN), Hieter P, Boycott KM, Campeau PM, & Bellen HJ (2017). Model Organisms Facilitate Rare Disease Diagnosis and Therapeutic Research. *Genetics*, 207(1), 9–27. 10.1534/genetics.117.203067 [PubMed: 28874452]
- Wohler E, Martin R, Griffith S, Rodrigues E. da S., Antonescu C, Posey JE, Coban-Akdemir Z, Jhangiani SN, Doheny KF, Lupski JR, Valle D, Hamosh A, & Sobreira N (2021). PhenoDB, GeneMatcher and VariantMatcher, tools for analysis and sharing of sequence data. *Orphanet Journal of Rare Diseases*, 16(1), 365. 10.1186/s13023-021-01916-z [PubMed: 34407837]

(a) Two-sided matchmaking



(b) One-sided matchmaking



(c) Zero-sided matchmaking

**FIGURE 1.**

Three types of genomic matchmaking to enable novel disease-gene relationship discoveries.

(a) Two-sided matchmaking refers to the scenario where two or more parties have the same novel candidate gene and are trying to find each other. This type of matchmaking is facilitated by the current MME. (b) One-sided matchmaking refers to the scenario where one party has a novel candidate variant or gene and queries a database housing structured genome-wide sequencing data from undiagnosed patients. (c) Zero-sided matchmaking refers to the scenario where computer algorithms are used to identify potentially matching undiagnosed cases with rare variants in the same gene and overlapping phenotypes.

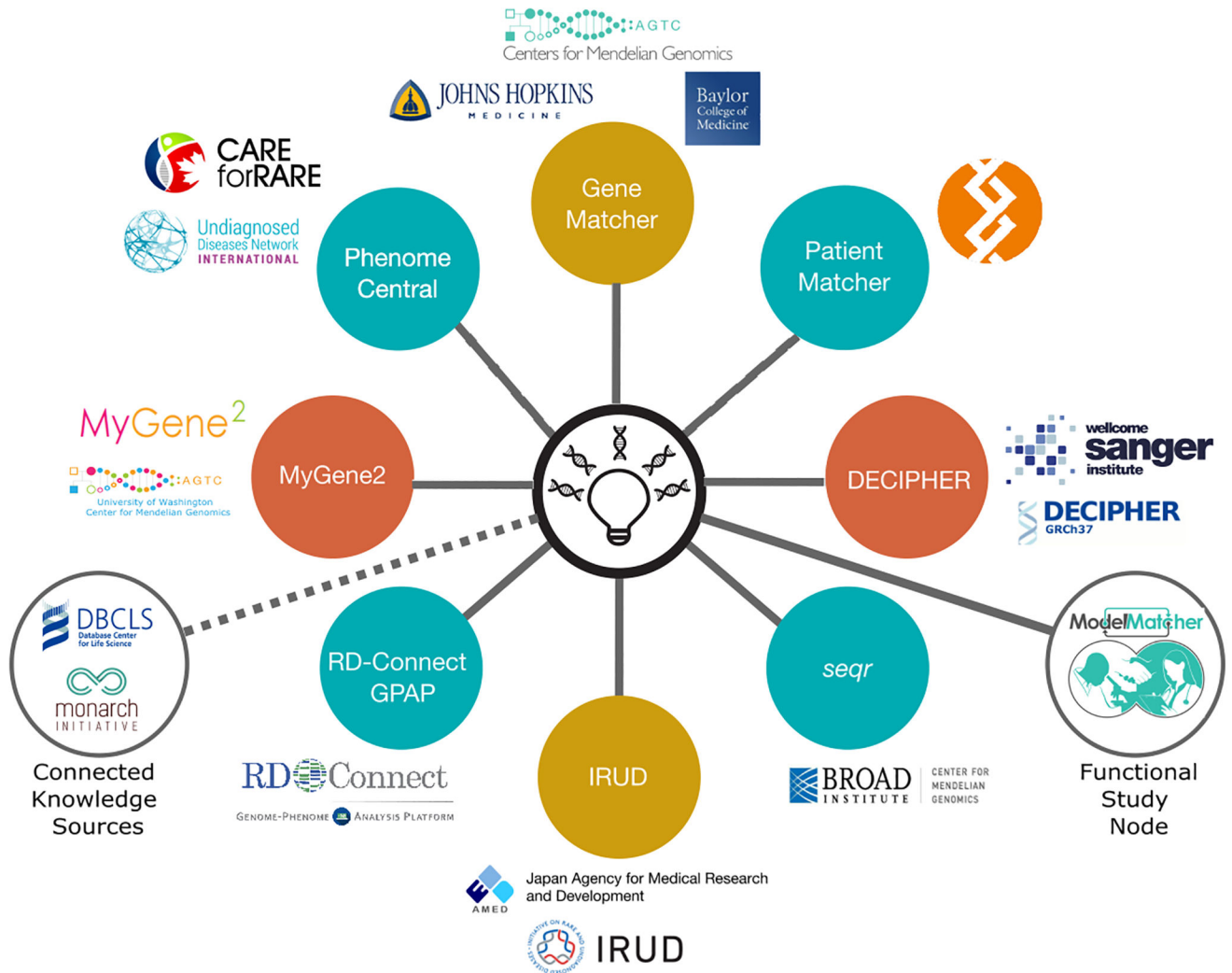


FIGURE 2.

The MME currently connects eight genomic matchmaking nodes (solid lines) representing >13,520 unique genes and 120,000 cases from over 12,000 contributors in 98 countries. The MME also connects to three databases that provide additional utility to the gene discovery scientific community and these databases are considered “Connected Knowledge Sources,” or “Model Organism Nodes”, specialized MME endpoints that go beyond the initial MME design of patient-patient matchmaking. The MME is a driver project of the GA4GH and the MME has been leveraging the expertise of the GA4GH working groups for guidance on pertinent aspects of the project.

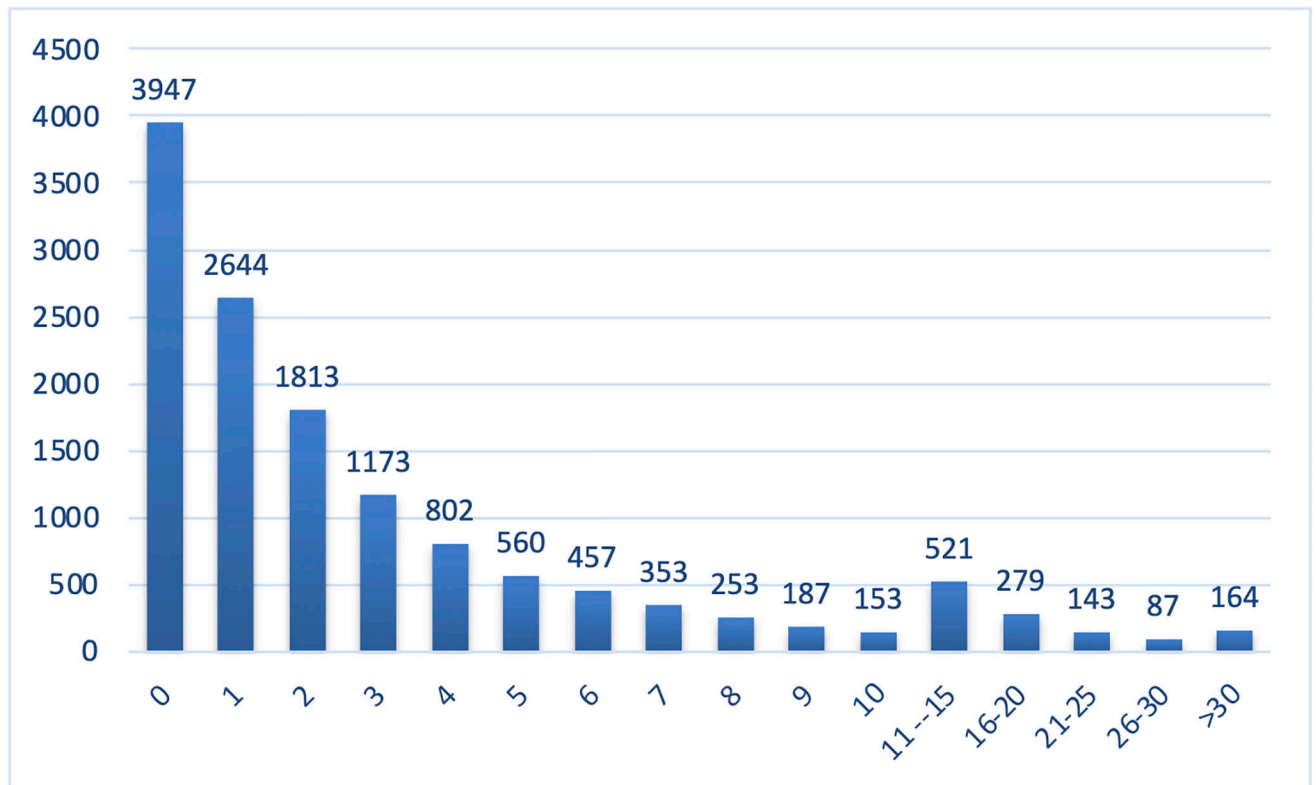


FIGURE 3.

Matches per gene in GeneMatcher. Among the 13,520 unique genes submitted to GeneMatcher as of March 1st, 2022, 3,947 (27%) have no match. The number of matches for genes with >30 matches ranges from 31 to 142 matches; about 0.1% (15) of unique genes in GeneMatcher have more than 60 matches.

TABLE 1

Details of the connected MME Nodes

Name	Location	Patients/Cases Total	In MME		
			Patients/Cases	Contributors/Submitters	Unique Genes
DECIPHER	United Kingdom	80,000	40,624	570	8,987
GeneMatcher	United States	62,718	62,492	11,031	13,520
IRUD	Japan	3,578	62	1	55
MyGene2	United States	4,492	1,600	304	1,313
PatientMatcher	Sweden	10,060	18	4	25
PhenomeCentral	Canada	12,775	8,974	444	3,027
RD-Connect GPAP	Europe	13,929	5,847	77	902
<i>seqr</i>	United States	7,929	1,193	84	1,242
Total		195,481	120,810	12,515	>13,520*

* The gene overlap across nodes has not been determined so the total genes is conservatively estimated using the largest single node.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

TABLE 2

Opportunities to improve the efficiency of the MME

Add additional details to submissions and information going across the MME (e.g., phenotypic features, variant details, zygosity of variants)
Remove or suspend submissions that are solved or add tags that a case is solved, but looking for additional cases for secondary genotype-phenotype publications
Remove or suspend submissions with candidates with very limited evidence
Generate collaboration templates for tracking cases, collaborators, and publications

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript