Care4Rare Canada: Outcomes from a decade of network science for rare disease gene discovery

Kym M. Boycott,[1,*] Taila Hartley,[1] Kristin D. Kernohan,[1] David A. Dyment,[1] Heather Howley,[1] A. Micheil Innes,[2] Francois P. Bernier,[2] Michael Brudno,[3] and Care4Rare Canada Consortium[4]

Summary

The past decade has witnessed a rapid evolution in rare disease (RD) research, fueled by the availability of genome-wide (exome and genome) sequencing. In 2011, as this transformative technology was introduced to the research community, the Care4Rare Canada Consortium was launched: initially as FORGE, followed by Care4Rare, and Care4Rare SOLVE. Over what amounted to three eras of diagnosis and discovery, the Care4Rare Consortium used exome sequencing and, more recently, genome and other 'omic technologies to identify the molecular cause of unsolved RDs. We achieved a diagnostic yield of 34% (623/1,806 of participating families), including the discovery of deleterious variants in 121 genes not previously associated with disease, and we continue to study candidate variants in novel genes for 145 families. The Consortium has made significant contributions to RD research, including development of platforms for data collection and sharing and instigating a Canadian network to catalyze functional characterization research of novel genes. The Consortium was instrumental to implementing genome-wide sequencing as a publicly funded test for RD diagnosis in Canada. Despite the successes of the past decade, the challenge of solving all RDs remains enormous, and the work is far from over. We must leverage clinical and 'omic data for secondary use, develop tools and policies to support safe data sharing, continue to explore the utility of new and emerging technologies, and optimize research protocols to delineate complex disease mechanisms. Successful approaches in each of these realms is required to offer diagnostic clarity to all families with RDs.

Introduction

With the genomic insights provided by the Human Genome Project and advances in high-throughput sequencing, we can now readily detect almost all DNA variation in a genome. One of the first areas of medicine to benefit from these developments was rare genetic disease (RD).[1] Given their rarity and often profound impact on reproductive fitness, thousands of RDs were largely intractable to the conventional discovery approaches used until a decade ago. The development of genome-wide sequencing (GWS) approaches, particularly exome sequencing, has now made the highly penetrant DNA-coding variants underpinning many RDs readily discoverable. Over the last decade, these developments have facilitated the identification of molecular causes of RD in the clinic, established new variant-disease associations in known genes (genes where deleterious variants are known to cause disease), and identified disease-causing variants in novel genes (genes with no previously known disease association) for these often devastating conditions.[1]

During this period, as the potential of GWS to identify new molecular causes of (or "solve") RDs became obvious, national and multinational gene-discovery projects emerged, including the UK Deciphering Developmental Disorders study (Health Innovation Challenge Fund and the Wellcome Trust Sanger Institute, 2011), the U.S. Centers for Mendelian Genomics (National Institutes of Health, 2012), and the EU NeurOmics and EURenOmics (European Commission, 2012). In Canada, we launched the Care4Rare Canada program, initially as Finding of Rare Disease Genes in Canada (FORGE), funded by Genome Canada and Canadian Institutes for Health Research (CIHR) in 2011. The success of these large-scale initiatives emphasized the importance of team science and relied on openness and willingness to share.[2–5] These approaches are particularly important given the nature of RDs; an individual clinician, researcher, or institution will not have sufficient experience, data, or resources to effectively identify disease-causing variants in novel genes on their own. After a decade of collaborative network science in Canada and three eras of RD gene discovery, the Care4Rare Canada Consortium recognized it was time to reflect on the lessons learned from our successes to best meet the challenges of RD diagnosis and discovery in the decade to come.

Care4Rare canada consortium

The Care4Rare Canada Consortium was launched in 2010 with a commitment to team science and open science as essential for the study of RD. Canada's geography and health system create challenges for studying RD: 38 million people spread across the world's second largest country receiving healthcare from 10 distinct provincial and 3 territorial publicly funded systems. Despite this, the tightly knit

[1]Children's Hospital of Eastern Ontario Research Institute, University of Ottawa, Ottawa, ON K1H 8L1, Canada; [2]Department of Medical Genetics and Alberta Children's Hospital Research Institute, University of Calgary, Calgary, AB T2N 4N1, Canada; [3]Department of Computer Science, University of Toronto, Toronto, ON M5S 2E4, Canada
[4]Further details can be found in the supplemental information
*Correspondence: kboycott@cheo.on.ca
https://doi.org/10.1016/j.ajhg.2022.10.002.

	Recruited in FORGE Era *(2011-2013)* Recognizable disorders and families with linkage analysis, Minimum of 1 Canadian family	Recruited in Care4Rare Era *(2013-2018)* N-of-1 families undiagnosed after standard-of-care testing (e.g., single-gene, gene panels, and microarray)	Recruited in C4R-SOLVE Era *(2018-2021)* N-of-1 families with non-diagnostic exome data generated by the healthcare system	Total Program
	375 Families	+830 Families	+601 Families	1806 Families
Outcomes from exome data				
Diagnoses in known genes	111 (30%)	263 (32%)	46 (8%)	420 (23%)
Diagnoses in novel genes	93 (25%)	71 (9%)	39 (6%)	203 (11%)
Percentage diagnosed	**204 (54%)**	**334 (40%)**	**85 (14%)**	**623 (34%)**
Compelling candidates	21 (6%); 3 VUS, 18 GUS	152 (18%); 88 VUS, 64 GUS	137 (23%); 74 VUS, 63 GUS	310 (17%); 165 VUS, 145 GUS
Discovery publications	63	40	18	121
Families with no candidates	150	344	379	933

Figure 1. A decade of rare disease (RD) gene discovery and diagnosis via the Care4Rare Canada Consortium
RD gene discoveries and diagnoses span three eras of the Care4Rare Canada program: FORGE, Care4Rare, and Care4Rare-SOLVE were a continuous series of large-scale pan-Canadian RD sequencing projects. For each era (colored boxes), the time frame, types of RDs studied, and number of families agreeing to participate is summarized. The outcomes, to date, from each era (gray box) display the diagnostic yield (in known and novel genes), compelling VUSs and GUSs, and novel genes published. Note that the outcomes (diagnosis or discovery) may have occurred at any point during the 10-year Care4Rare Canada program. A list of our candidate GUSs can be found in the Open Access Data webpage of the Genomics4RD website (https://www.genomics4rd.ca/openaccess). Families whose RD remains unsolved following clinical exome sequencing are eligible to participate in our research protocol for unsolved RDs. VUS, variant of uncertain significance; GUS, gene of uncertain significance.

Canadian genetics community (~120 clinical geneticists working from 21 clinical genetic centers across the country) quickly came together when Genome Canada launched the Advancing Technology Innovation through Discovery Program in partnership with CIHR. The first of Care4Rare program's phases, FORGE, was funded through this program to test the utility of exome sequencing for identifying the genetic causes of childhood diseases and operated from 2011 to 2013. The success of this pilot led to the second phase, Enhanced CARE for RARE Genetic Diseases in Canada (Care4Rare; 2013–2018), and third phase, Care4Rare Canada: Harnessing Multi-omics to Deliver Innovative Diagnostic Care for Rare Genetic Diseases in Canada (Care4Rare-SOLVE; 2018–2023), with funding from Genome Canada, CIHR, and other partners. Each of these studies was designed to address the challenges of the previous phase, and participating families remain in the Care4Rare program until their RD is solved.

The Care4Rare Canada Consortium's six informal guiding principles—support meaningful work, empower and engage front-line clinicians, build capacity at all levels, foster respectful collaborations, share generously, and partner with the RD patient community—worked particularly well for RD in Canada given the familiarity and shared history of our clinical, scientific, and RD communities. To date, the Care4Rare Canada Consortium has studied 1,806 families with suspected RD (see Figure 1 and eligibility criteria outlined in Table S1). Our program has provided molecular diagnoses to 623 families (34%) who had completed standard-of-care genetic investigation prior to enrollment. Of these diagnoses, 420 (23% of participating families) were in genes where deleterious variants are known to cause disease, and 203 (11% of participating families) were in genes where deleterious variants had not yet been associated with disease, with 121 novel gene discoveries published to date (Table S2); the remainder are in various stages of confirmation studies and manuscript preparation. We have also identified compelling candidates for 310 families (17%): 165 variants of uncertain significance (VUSs) and 145 genes of uncertain significance (GUSs) that are undergoing further study (Figure 1). In this perspective, we describe a decade of Canadian experience with RD diagnosis and gene discovery. For each era (FORGE, Care4Rare, and Care4Rare-SOLVE), we describe our program outcomes (Figure S1), challenges we encountered and tools we developed to address them, clinical impacts arising from our Consortium (Figure S1), and key lessons learned.

The three eras of the Care4Rare canada consortium

FORGE era (2011–2013): A golden age of exome-based gene discovery for recognizable syndromes

The FORGE Canada project was launched as a two-year pilot in April 2011 to rapidly identify molecular diagnoses associated with a wide spectrum of rare pediatric-onset single-gene disorders for which at least one Canadian patient was known.[2] As a starting point, our GE^{3}LS (Genomics and its Ethical, Environmental, Economic, Legal, and Social

Aspects) team developed study documents, a model consent form based on best practices related to GWS for RD families (including the necessary disclosure of research results back to families and their care teams), and protocols for the sharing of data and biological samples within Canada and internationally.[3] These materials received research ethics approval at each of the 21 participating Canadian sites, which effectively harmonized Canada's community and collaborative process for RD genomic research.

Following a national call, clinical geneticists across Canada submitted hundreds of proposals for RDs to be studied during FORGE, which were reviewed by our pan-Canadian steering committee. The committee selected 264 RDs, which were rapidly assembled for study through samples contributed by the Consortium and collaborations with clinicians from 17 different countries and which represented 375 unique families (1,000 participants from Canada and 300 international participants). The FORGE project used exome sequencing and employed Sanger sequencing and functional assays on human cell lines for secondary validation purposes. RDs studied were highly enriched for a single-gene etiology: e.g., two-thirds of disorders studied were either cohorts of highly recognizable syndromes, or they exhibited significant family histories with recurrence of a disorder and/or known consanguinity or were from isolated communities likely to have founder mutations.

The FORGE era was significant for the number of discoveries made related to recognizable syndromes, including causative genes identified for Hajdu-Cheney (MIM: 102500),[6] Nager (MIM: 154400),[7] Floating-Harbor (MIM: 136140),[8] Weaver (MIM: 277590),[9] megalencephaly capillary malformation (MIM: 602501),[10] Chudley-McCullough (MIM: 604213),[11] mandibulofacial dysostosis with microcephaly (MIM: 610536),[12] SHORT (MIM: 269880),[13] two types of French Canadian Joubert syndrome (MIM: 614615, 614970),[14,15] and microcephaly-capillary malformation syndrome (MIM: 614261).[16] At the time of publication, we have identified one or more molecular diagnoses for 54% (204 of 375) of the original FORGE families (Figure 1). This includes 93 families with molecular diagnoses in novel genes, from which 63 publications have been the primary studies describing disease-causing variants in the novel gene (Table S2). In addition, we have identified compelling candidates in 21 (6%) of 375 FORGE families, including 3 VUSs and 18 GUSs.

Care4Rare era (2013–2018): Solving N-of-1 RDs

The Care4Rare Canada era was a 5-year RD diagnosis and discovery project built on the success of FORGE, which continued to leverage the Consortium membership and our pan-Canadian Research Ethics Board (REB)-approved GWS research protocol. Care4Rare went beyond FORGE to study a wider range of pediatric and adult-onset unsolved RDs. An additional 830 families agreed to participate, and more than 2,000 participant samples were obtained during this era. Care4Rare primarily used exome sequencing but also conducted small pilots of genome sequencing and RNA sequencing to gain experience with the emerging technologies of this era.

From the Care4Rare era, we identified one or more molecular diagnoses for 334 of the 830 (40%) families (Figure 1). Like during FORGE, GWS was not part of the clinical standard-of-care for this cohort, and thus there was a high proportion of families with molecular diagnoses in known genes (263 of 830, 32%). Another 71 families (9%) received molecular diagnoses representing novel gene discoveries, including 40 for which we led or contributed significantly to the discovery publication (Table S2). We continue to study the 152 families (18%) for whom we identified compelling candidates, including 88 VUSs and 64 GUSs. This high proportion of GUSs represents mostly N-of-1 findings (only one family with a particular GUS in the cohort) and reflects the challenges of a large number of single, small families from this era for which we have insufficient evidence for disease causality. In this era, we discovered only a handful of disease-causing variants in novel genes causative for highly recognizable syndromes, such as cerebrocosto-mandibular syndrome (MIM: 117650)[17] and encephalocraniocutaneous lipomatosis (MIM: 613001),[18] and deleterious variants in novel genes associated with genetically heterogeneous RDs like French Canadian Joubert syndrome (MIM: 616781).[19] This low number reflects the relative rarity of recognizable syndromes that remained unsolved with sufficient samples available for study. In addition, Care4Rare's research mandate extended beyond discovery to delineate the genotypic and phenotypic spectrum of recognizable RDs that we previously discovered, such as Floating-Harbor syndrome and mandibulofacial dysostosis with microcephaly.[20,21] We also described functional insights into disease mechanisms and delineated the diagnostic utility of GWS in different RD cohorts,[22–24] including the presence of multiple genetic diagnoses in 3.5% of probands studied (Box S1).[23]

SOLVE era (2018–2023): Maximizing the exome to prioritize families for multi-omics

The Care4Rare-SOLVE project is a five-year project focused on providing a diagnosis for families with suspected RD following uninformative exome sequencing (clinical or research) and facilitating access to clinical GWS for all Canadians. SOLVE uses a stepwise approach that incorporates data-sharing to diagnose RDs: first using reanalysis to maximize the discovery potential of the exome, followed by a discovery pipeline of other 'omic technologies and approaches (genome sequencing, RNA sequencing, deep sequencing; see "A research protocol for unsolved RD").[25,26] The SOLVE cohort includes 667 families from the FORGE and Care4Rare projects whose RD remained unsolved at the start of SOLVE (April 2018), as well as an additional 1,104 families (representing 3,254 participants) who have consented to participate during SOLVE following non-diagnostic exome sequencing.

The final results of the SOLVE project remain to be determined, as we continue to analyze data and invite families

to participate. As of the end of 2021, we have completed the exome reanalysis for 601 of the 1,104 new families enrolled during SOLVE and identified one or more molecular diagnoses for 85 (14%) (Figure 1). Unsurprisingly, given that most of these families already had clinical exome sequencing testing, there were fewer diagnoses in known genes (46 of 601, 8%), although this remains a significant yield, and disease-causing variants in novel genes reported since the clinical analysis was the predominant contributor. Another 39 families (6%) have received molecular diagnoses in novel genes, including 18 for which we led or contributed significantly to the discovery publication (Table S2). Compelling candidates were identified for 137 families (23%), including 74 VUSs and 63 GUSs. SOLVE has also focused on a few of the remaining unsolved yet recognizable syndromes (e.g., PHACES [MIM: 606519], Aicardi syndrome [MIM: 304050], Hallermann-Streiff syndrome [MIM: 234100], and Dubowitz syndrome [MIM: 223370]), of which only Dubowitz syndrome has been resolved to be secondary to extensive locus heterogeneity,[27] which resulted in removal of this condition from the most recent edition of Smith's Recognizable Patterns of Human Malformations.[28]

Many families without a compelling candidate variant over the past decade were selected for short- or long-read genome sequencing, RNA sequencing, and/or deep sequencing, and analyses are ongoing (Figure 1); our research protocol for unsolved RDs is described under Strategies. Thus far, this approach has been triggered for 235 families, including 31 families recruited originally in the FORGE era, 50 families recruited in the Care4Rare era, and 154 families recruited in the SOLVE era.

Strategies to address challenges across the decade

The pace of our discoveries and types of RDs we solved evolved over the decade and were influenced by the families we studied, advances in genome and information technology, and available data-sharing solutions. As comprehensive gene panels and exome sequencing were integrated in the clinical diagnostic pathway for RD, molecular diagnoses made in known disease genes decreased by two-thirds: accounting for only 8% of SOLVE diagnoses compared to 30% for FORGE. Our pace of novel gene discovery slowed as RDs become more challenging to solve: our FORGE discoveries were from large cohorts with recognizable syndromes and a single causative gene, whereas we now tackle N-of-1 RDs with complex genetic mechanisms outside the reach of the exome (Figure 2). Through a decade of RD discovery, we approached key bottlenecks and challenges with the infrastructure needed to address them.

Standardizing the capture of phenotype

During FORGE, a clinical diagnosis was sufficient for study eligibility and GWS interpretation; however, as we began to study less-recognizable diseases or diseases without a name, the need for standardized capture of phenotypic descriptions became apparent. Thus, toward the end of FORGE, we developed the PhenoTips software to support the routine collection of computer-readable phenotypic data using human-phenotype ontology (HPO) terms.[29] PhenoTips continues to be used by the Consortium to this day and is the first step to entry into the Care4Rare discovery pipeline. PhenoTips has been integrated into the Epic electronic health record at several Canadian institutions and embedded into clinical genetics workflows for GWS-based diagnostic testing, facilitating the transfer of phenotypic data from affected individuals who provide informed consent for research.

Matchmaking candidate genes with the international community

The next central challenge we faced was how to efficiently generate evidence for causality of a GUS linked to a N-of-1 RD, which were not clinically recognizable and often didn't have a name. While traditional networking (i.e., personal contact via email or conferences) continues to be helpful (Figure 2), it is not scalable. Care4Rare applied two approaches to address this challenge: global data-sharing to identify affected individuals with an overlapping phenotype and compelling variants in the same gene, and catalyzing connections with the Canadian model organism community (see below). We launched PhenomeCentral in 2014 as a data-sharing solution that acts as a centralized repository for unsolved RDs[30,31] and contains the complete set of phenotypic, genotypic, and candidate gene data from a decade of the Care4Rare program, as well as data from Undiagnosed Diseases Network in the U.S. and other groups from around the world. However, early into the Care4Rare project, we realized that there were several other RD datasets containing data to similar PhenomeCentral that existed in siloes. Thus, in collaboration with the International Rare Disease Research Consortium (IRDRIC) and the Global Alliance for Genomics and Health (GA4GH), we engaged with key stakeholders to collectively launch the Matchmaker Exchange (MME) in 2015 (Figure 2).[32]

MME facilitates two-sided matchmaking, where two or more parties try to identify others with the same novel candidate gene where disease-causing variants have not yet been reported.[32] Operating as a federated network connecting databases of candidate genes and/or genotypes and rare phenotypes using a common application programming interface,[32] MME currently connects eight databases including, among others, PhenomeCentral,[31] GeneMatcher,[33] DECIPHER,[34] RD-Connect GPAP,[35] and seqr.[36] With more than 120,000 submissions and 13,000 unique genes shared across MME from over 12,000 submitters from 98 countries, matchmaking via MME has become central to gene discovery,[37] providing diagnostic clarity for families around the world. Care4Rare credits MME for facilitating the discovery and publication of disease-causing variants in 26 novel genes during the last seven years

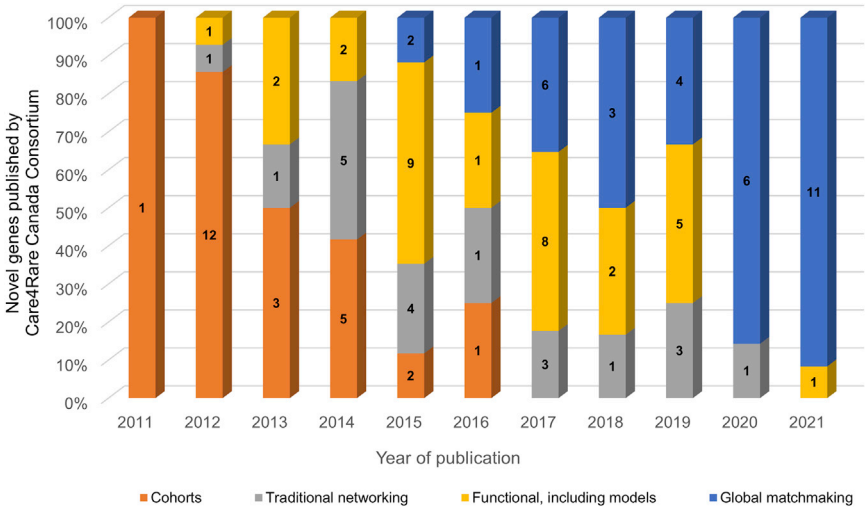

Figure 2. Novel genes published by Care4Rare Canada, by year of publication and primary approach to support of disease causality

Gene discoveries in the early days of Care4Rare Canada relied on analyzing single-RD cohorts (orange) or identifying similarly affected individuals through traditional networking via email or conferences (gray). Starting in 2014, following the launch of the Canadian Rare Diseases: Models and Mechanisms (RDMM) Network, functional assays and model organisms (yellow) became an important approach to providing supporting evidence for disease causality. PhenoTips to capture clinical data as HPO terms came online in 2013 and facilitated the development of PhenomeCentral, a matchmaking data platform, in 2014. Following the launch of Matchmaker Exchange (MME) in 2015, with PhenomeCentral as one of the original three nodes connected, traditional networking was replaced with automated global matchmaking (blue) with researchers beyond our usual collaborations. MME-fueled gene discoveries continue to be the dominant discovery approach. While Care4Rare discovered disease-causing variants in 121 novel genes (listed in Table S2), Figure 2 includes only the 109 novel genes that represent the primary discovery using the described approaches. Of the 12 novel genes not included, seven were the second publication, two used a pilot matchmaking algorithm not yet released, and three were case reports with compelling genetic evidence.

(see references for examples[38–47]) (Figure 2 and Box S1) and have shown that the MME facilitates collaborations that lead to evidence for pathogenicity and publication for 15% of novel candidate genes submitted.[48] Our analyses have highlighted both the success as well as the labor-intensive nature of the MME, leading to publication of recommendations for improving the specificity and efficiency of the MME.[37]

Collaborating with the model organism community

Model organisms can provide important evidence to support the pathogenicity of deleterious-appearing variants in novel gene candidates. However, early into the Care4Rare era we realized that it was difficult for clinician-investigators studying RDs to do this work: the model organism community was largely unknown to them and there was limited funding for such research. In response, we co-launched the Canadian Rare Diseases: Models and Mechanisms (RDMM) Network in October 2014. RDMM is a Canadian consortium that connects clinicians discovering the molecular mechanisms of RDs with more than 680 Canadian model organism scientists with expertise in more than 8,800 genes.[49] RDMM uses a committee process to identify and evaluate collaborations between clinicians and Canadian model-organism scientists and provides 25,000 Canadian dollars in catalyst funding for projects to evaluate pathogenicity and begin to elucidate the molecular mechanisms of RD. For example, matchmaking via MME was unsuccessful for *CEP55* (MIM: 610000), for which a homozygous truncating variant was identified in a single family with three affected individuals with MARCH syndrome (multinucleated neurons, anhydramnios, renal dysplasia, cerebellar hypoplasia, and hydranencephaly; MIM: 236500). CRISPR-Cas9-mediated ablation of *CEP55* in zebrafish embryos recapitulated the phenotypic features of MARCH syndrome and provided functional evidence of the pathogenicity of the homozygous loss-of-function variant.[50] This subsequently led to several families being identified with this RD.[51] RDMM continues to catalyze and support early work on novel genes discovered by Care4Rare, with funding in place until 2027. This approach was subsequently modelled in Europe, Japan, and Australia, who are all co-founding partners of RDMM International along with the Canadian RDMM.

A RD data lake for Canada

While matchmaking via MME revealed the importance of global data sharing to address the N-of-1 challenge, solving RD requires additional solutions that can share different levels and types of data. Thus, we made significant investments during the SOLVE era to facilitate the safe and responsible sharing of diverse sets of data from Canadian research participants with RD. In February 2021, we launched Genomics4RD to centralize and harmonize, in a retrospective and prospective manner, deep phenotypic and multi-omic Canadian RD data, as well as the results of diverse investigations and patient and health system outcomes.[52] Genomics4RD is a web-accessible platform designed to share data nationally and internationally for RD gene discovery, diagnosis, and other related research.[52] Data-use restrictions or conditions are attributed to datasets using the GA4GH's Automatable Discovery and Access Matrix (ADA-M) codes to ensure data are used according to its informed consent.[53] Genomics4RD provides three levels of data access: open (e.g., public access to lists of candidate genes), registered (RD researchers outside of Canada for specific use cases), and controlled (researcher with an REB-approved protocol with full access to specific data).

To support analyses across the Consortium, Genomics4RD offers a "variant store": a user interface to efficiently query for variants by genomic position, gene name, and predicted pathogenicity, as well as frequency in population and control datasets and our own cohort.[52] Data depositors can access the larger Genomics4RD dataset to explore genes of interest within the platform and identify participants that have the same RD (Box S2). We retrospectively harmonized and transferred the complete dataset from the FORGE and Care4Rare projects to Genomics4RD, while all SOLVE data are prospectively captured. Genomics4RD has been integrated with PhenomeCentral for two-sided matchmaking via MME. In addition, we are piloting a one-sided matchmaking platform for registered access users: this feature allows users to query a gene of interest, and Genomics4RD returns a list of variants associated with the gene, along with variant- and case-specific details, and provides the ability to contact associated data depositors. We will continue to develop additional tools and data-sharing strategies for the diverse data types available in Genomics4RD to facilitate RD discovery and, ultimately, diagnoses for individuals affected with RD and their families.

A research protocol for unsolved RD

While our early discovery pipeline was successful, RDs remaining unsolved are increasingly complex and challenging to resolve. SOLVE is tackling discovery beyond the nondiagnostic exome and begins with reanalysis of existing exome sequencing data as the entry into the protocol. With SOLVE, we are exploring the diagnostic utility of genome sequencing (short- and long-read), transcriptome sequencing, deep sequencing, metabolomics, lipidomics, and epigenomics. To support this work, we developed a research protocol that matches a family with an unsolved RD to the most informative set of technologies based on our hypothesis as to why their RD remains unsolved.[26] We are evaluating this protocol starting with 235 families who remain without a diagnosis, selected from the three eras of our program. Using this protocol, we categorize each family into one of four groups based on their clinical presentation (Group 1: limited genetic heterogeneity, Group 2: unsolved recognizable RD, Group 3: high degree of genetic heterogeneity, Group 4: RD without a name) and then begin a group-specific research pipeline of technologies that address the reason(s) their RD likely remains unsolved (Table S3).[26] Our preliminary data for Groups 3 and 4 shows us that 40% and 25%, respectively, of these families are being solved by the technologies used thus far, but additional analysis is ongoing, and the goal will be the development of best research practices for these groups of individuals.

Impact of the Care4Rare program

Building new clinical genomic knowledge

Definitively establishing the pathogenicity of a variant, or set of variants, for an RD requires robust evidence, as does associating deleterious variants in that gene with an RD for the first time. These evidence standards evolved over the decade, with new guidelines published for variant interpretation and thresholds for evidence for novel gene discovery.[54,55] The Care4Rare Consortium is committed to the timely translation of our robust discoveries to clinical care and patient benefits. Over three eras of diagnosis and discovery, we have shared hundreds of variants in known disease genes with the international community through publication or deposition in knowledge bases (e.g., ClinVar[56] and LOVD[57]). Thus far, Care4Rare has contributed 121 publications describing disease-causing variants in novel genes, with another 28 in preparation. We have also identified 145 compelling GUSs: they are being shared via MME (to identify additional affected individuals), and a subset are being studied using patient cell lines or via the RDMM network. The list of candidate genes is available for viewing on the Genomics4RD website.

Facilitating access to clinical GWS for Canadians affected with a RD

The vision of many health research programs is to implement their innovation into the healthcare system to improve patient care. FORGE demonstrated, for Canada, that exome sequencing could be used to discover new genetic causes of RDs as well as provide a molecular diagnosis for affected individuals with a known but undiagnosed RD (30%; Figure 1). The results called for a proposal, from the Canadian College of Medical Geneticists (CCMG), for Canadian recommendations regarding the use of GWS for RD diagnosis in clinical practice.[58] It also contributed to the international debate at the time regarding how best to manage incidental findings, highlighting parents' enthusiasm for returning such results from genetic research to children.[59] During Care4Rare, we demonstrated that GWS could provide a molecular diagnosis in known genes for 29% of individuals who completed standard genetic investigation, highlighting the utility of this approach at the end of the diagnostic odyssey.[24] We led the development of the CCMG position statement on the clinical application of GWS for monogenic disease,[60] which paved the way for clinical adoption across Canada. The implementation of GWS was further supported by our work measuring the impact and value of a diagnosis for the well-being of families with unsolved RD.[61,62]

Our experiences during the FORGE and Care4Rare projects led us to the realization that deeper engagement with healthcare was needed. For SOLVE, we co-designed with the ministries of health in Alberta and Ontario an evaluation of the diagnostic utility, clinical utility, and value for money of publicly funded exome sequencing for unsolved RD. Informed by our preliminary data from 250 participants, we co-designed with policymakers a clinical implementation pathway for new 'omic technologies. As a result, pilot implementation of clinical GWS for RD diagnosis, with GWS performed in public diagnostic laboratories (for example see Hayeems et al., 2022[63]) is now happening

across the country. Collectively, the provinces are working to build a Canadian data ecosystem that will support high-quality clinical GWS and provide research opportunities for families with RD. Further, as GWS has become standard-of-care for the diagnosis of RD in children and adults, there is interest in performing this test in the prenatal setting. Analogous to our work in the postnatal setting, we co-led the development of a CCMG position statement on the clinical application of fetal GWS for the investigation of multiple fetal anomalies.[64] We will continue to use our clinical implementation pathway to integrate other 'omic technologies into the healthcare system.

Understanding RDs in some of Canada's unique populations

Canada's population of approximately 38 million includes diverse ethnicities and cultures and includes a number of founder populations. Over the past ten years, we have had the privilege to collaborate with several uniquely Canadian populations, providing those communities with information about RDs caused by a founder effect. In some instances, we identified a founder pathogenic variant in a known disease gene (e.g., *SLC34A3* [MIM: 609826] and hypophosphatemic rickets with calcinuria in the Hutterite population [MIM: 241530][65]) or a phenotypic expansion of a known disease (e.g., *PRUNE1* [MIM: 617413] and a childhood neurodegenerative disease in Manitoba Cree families [MIM: 617481][66]). Such discoveries provided important data for early diagnosis and, in some instances, led to newborn- and carrier-screening programs. We also delineated the significant contribution of some genes in causing a genetically heterogeneous disease presentation in a population, for example, the spectrum of *SPG7* (MIM: 602783) mutations in French-Canadian spastic ataxia (MIM: 607259).[67] In other instances, we discovered a novel gene that is now recognized to cause RD in affected individuals around the world, for example *GPSM2* (MIM: 609245) causing Chudley-McCullough syndrome and *CEP55* causing MARCH syndrome in the Mennonite population,[11,50] *SGPL1* (MIM: 603729) causing steroid-resistant nephrotic syndrome in the Hutterite population (MIM: 617575),[68] and multiple genes causing Joubert syndrome in the French Canadian population (*C5ORF42* [MIM: 614571],[14] *TMEM231* [MIM: 614949],[15] and *CEP104* [MIM: 616690][19]). Finally, there is that rare gene discovery that provides insight into therapy: the identification of a homozygous pathogenic variant in *SLC39A8* (MIM: 608732) causing a neurodegenerative syndrome secondary to manganese deficiency in the Hutterite population led to early manganese supplementation for affected infants (MIM: 616721).[69]

Education

Alongside RD discovery and diagnosis, the Care4Rare program supports a multi-pronged approach to clinical translation. Thus, we have invested in training and education to build clinical capacity for interpreting 'omic data and further support clinical implementation. To build capacity in RD genomics, we developed a "Genomics Technologies for Patient Care" program for CCMG laboratory trainees across Canada. This program consists of four modules (Technical Aspects, Ethical Issues, Clinical Reporting, and Clinical Cases), which has been incorporated into the CCMG's routine training program. We also worked with the Royal College of Physicians and Surgeons of Canada to build capacity in their specialty training program (residency) in Medical Genetics and Genomics. In some Canadian provinces, residents in this program spend a mandatory month with their local Care4Rare team to interpret GWS for discovery. We developed a clinical fellowship program, "Genomic Medicine for Rare Diseases," which is offered yearly. We also offer 2- to 6-week electives to residents, fellows, genetic counseling students, and staff across Canada to work with the Care4Rare team to gain experience in the interpretation of 'omic data. Finally, we delivered workshops for trainees and faculty to provide hands-on training in analyzing and interpreting exome sequencing data. The training of the next generation of providers and upskilling the current generation are important approaches towards providing a diagnosis for all families with RD.

Advocacy and engagement

The Care4Rare program has benefited greatly from our partnership with the Canadian Organization for Rare Disorders (CORD). With their support, we conducted a discrete-choice experiment survey among parents of children with RD to measure the value of diagnostic testing; parents were willing to pay CAD$6,590 for exome sequencing,[62] and these data supported advocacy with provincial ministries of health to fund GWS programs. Importantly, we engaged with CORD to contribute to Canada's first rare disease strategy: "Now is the Time: A Strategy for Rare Diseases is a Strategy for all Canadians."[70] To shine a light on the costs associated with RD in Canada, we performed a retrospective cohort study using population-based provincial health administrative data for children with RD in Ontario. We measured direct healthcare costs and resource use 5 years after a diagnosis.[71] Costs were 4.5–19.8 times higher for the RD cohort compared to the general (control) population or to asthma or diabetes cohorts. This work sparked considerable interest in the socioeconomic burden of RD in Canada and led to additional funding opportunities from CIHR. Continued partnership with community stakeholders will ensure the impact of our work going forward.

Lessons learned

Identify a productive team and establish a supportive environment

Owing to thousands of individually rare diseases, RD research demands team and interdisciplinary approaches to discovery. Expertise in clinical and molecular genetics, informatics, computer science, biology, health systems, health economics, ethics, and policy, to name just a few,

are necessary. Care4Rare has always represented a coalition of the willing. Understanding the team and how they work and recognizing that the team's membership will be dynamic have been key to our success. Essential for broad participation is a strong, supportive, and responsive coordinating center to meet the needs of the network and do the heavy lifting. Recognition of contributions and promoting and supporting the leaders of each discovery has ensured engagement across the country. Finally, and most importantly, keeping participating families at the center of everything we do ensures that any issues that might rarely arise within a discovery team are put in the appropriate context and resolved as quickly as possible.

Clear eligibility criteria are needed when resources are limited

Over the past decade, Care4Rare has spent a great deal of time ensuring that the families studied in our discovery pipelines are quite likely to have a RD with a genetic etiology. At times, this has meant we have had to make tough choices and establish clear eligibility criteria regarding selection of participants. Though not as significant of a challenge in Canada, we also had to ensure that the healthcare that the participating family underwent in the system met standard-of-care to minimize research dollars being spent on investigations that should have funded by the relevant ministry of health. Deeply phenotyped participants, along with family members, that meet tight clinical criteria ensure that our research dollars are spent in the best possible fashion.

Broad data sharing is critical

We realized early that traditional networking and functional studies would not compensate for the lack of large cohorts of similarly affected individuals. Global data sharing is critical to identify RD cohorts and establish causation. Global matchmaking, via MME, was developed to address the N-of-1 challenge, and this has been essential for gene discovery over the past 7 years, particularly for programs the size of Care4Rare. We recognize that MME has been a critical tool and a leading source of RD gene discoveries since 2019 (Figure 2), but it is resource intensive.[48] Upon reflection, the mandatory contribution of phenotype and inheritance pattern for all databases connected via MME would have made global matchmaking much more efficient. The need for even larger cohorts, enabled by efficient global matchmaking, has become evident as we begin to analyze more complex 'omic data for families that remain without a compelling coding variant after exome sequencing.

Engage policymakers early

As the Care4Rare Consortium began to recognize the diagnostic potential of GWS for families with suspected RD, implementing exome sequencing into clinical practice became a priority. In the Care4Rare project, we set out to generate the evidence that we thought policymakers and payers needed. We realized the error with this approach as we began working alongside policymakers and payers in unrelated projects. Rather than starting with the development of researcher-driven evidence, policymakers should be actively engaged at the outset so research is co-designed with them to meet their needs. As such, as part of the SOLVE project, we collaborated with ministries of health to co-design a study evaluating the utility and value of GWS for RD diagnosis. As a result, we were able to integrate GWS into the clinic while developing an evaluation framework to inform future decisions about the public funding of 'omic technologies for RD diagnosis.

It is only going to get tougher to make novel discoveries

In our reflection on the last decade, we would be remiss if we did not remark on the increasing effort required for discoveries in each era. In the initial FORGE era, exome discoveries often leapt from our spreadsheets, to the excitement of our team. The following N-of-1 era brought a mix of projects with incredibly compelling exome candidates and others that, despite convincing phenotypes and/or family history, did not have any interesting findings. Herein the effort required for each discovery began to increase: projects with compelling candidates required laboratory functional studies and matchmaking, both of which are highly effective in providing necessary evidence for causality, but are time consuming. In our current multi-omics era, each project has required a further and substantial increase in analysis time, taking weeks to sift through data and often ending in disappointment. Deciphering the complexity of multi-omic data to identify RD disease mechanisms is an ongoing work in progress and requires large amounts of time and resources, which is unlikely to change in the near future.

The next era

Through the efforts of the Care4Rare Canada Consortium and similar initiatives around the world, GWS has become standard-of-care for the diagnosis of RD in many jurisdictions. GWS has lived up to the promise of a transformative technology, with an outstanding diagnostic yield of approximately 30%, depending on the inclusion/exclusion criteria used. As GWS is integrated into the diagnostic care pathway for RD, research programs will need to shift from generating GWS data to accessing it. Leveraging clinical GWS data from families with solved and unsolved RDs will be critical to RD research, sustaining the gene discoveries that fuel clinical diagnosis. To do so, we must develop systematic approaches for secondary use of data generated in the diagnostic setting. However, clinical data are under the protection of a variety of healthcare administrations, and datasets are siloed within a variety of data custodians. We have found that health systems have not kept pace with families' desire to share their data, and access to RD research is not equitable

across different jurisdictions. Thus, with funding from Genome Canada's All for One precision health program, we launched an 18-month project to define a Canadian health data ecosystem for RDs. We are exploring a two-pronged data solution: Knowledge Network is a pan-Canadian variant database that will facilitate high-quality clinical GWS, while Connect is a patient research registry that will provide equitable access to precision health research for Canadians with RD. Incorporating consent-to-recontact language, initially as part of all GWS testing but expanding to any genetic testing in Canada, would provide a route to sharing data and providing equitable research opportunities for all Canadians with RD.

Care4Rare, along with many international groups, has demonstrated the strengths and opportunities for further efficiencies of the current global genomic matchmaking platform, MME.[37] The next era will see advances in genomic matchmaking approaches, using more automated protocols alongside better algorithms, required for matchmaking on a global scale. Initiatives like the GA4GH,[72] the IRDiRC,[73] and the World Economic Forum's Breaking Barriers to Health Data project[74] will continue to promote and support data-sharing initiatives for genomic and RD research. Data-sharing solutions that accommodate a variety of 'omic data are emerging, but mechanisms and policies for sharing this type of data between databases must be developed. Furthermore, data sharing will need to expand beyond 'omic data and HPO terms to capture deeper quantitative and unstructured clinical data.

The next era of RD diagnosis and discovery will focus on very-difficult-to-solve RDs, those for which the cause likely lies outside or is difficult to detect within the coding genome, is oligo or polygenic in nature, or is secondary to more complex epigenetic or gene/environment interactions. The biological basis of penetrance and expressivity will need to be explored. The remaining recognizable diseases that have not yet been solved bring their own set of data-sharing challenges.[75] It is likely that different types of samples will be required (e.g., affected tissue or tissue during specific developmental time point) to solve these RDs, and sharing these important resources will be essential. Best practices on the research approaches to these unsolved diseases will ultimately need to be developed and shared.

Finally, it is important that the stakeholders (e.g., healthcare systems, policymakers, and funders) not lose focus on the goal to understand the molecular mechanism of all RDs, as there is still much they have to teach us about both biology and health. Prioritizing funding, open science, and data sharing will ultimately benefit the individuals and families living with RD.

Conclusion

Over the last decade, the Care4Rare Canada Consortium has investigated the molecular cause of RD for 1,806 families. Although we have provided a diagnosis for many of these families, we continue to study the most challenging RDs for those that remain unsolved after exome sequencing. Over this period, we have adapted and expanded our approaches as the landscape of unsolved RDs has changed: moving beyond recruitment of recognizable diseases, sequencing, and analysis to global data sharing and utilization of multiple different 'omics to tackle complex disease mechanisms. The challenge of solving all RDs remains enormous, and the work is far from over. We must leverage clinical genomic and other 'omic data for secondary use, continue to develop tools and policies to support safe data sharing, and optimize research protocols to delineate complex disease mechanisms. Successful approaches in each of these realms are required to offer diagnostic clarity for all families with RD.

Supplemental information

Supplemental information can be found online at https://doi.org/10.1016/j.ajhg.2022.10.002.

Acknowledgments

The authors would like to thank the Canadian clinical and scientific RD communities for their strong collaborative spirit. We are grateful to the participants and their families, the Canadian Organization for Rare Disorders, Dennis Bulman, Alex MacKenzie, Cindy Bell, Marc LePage, Paul Lasko, Martin Osmond, and Kevin Keohane for their valued partnership and support over the past decade. We thank the Genome Canada technology platforms: The Centre for Applied Genomics (TCAG), McGill Applied Genomics Innovation Core (MAGIC), and the BC Cancer Agency Genome Sciences Centre Genomics Technology Platform. We would also like to thank the Centre for Health Genomics and Informatics at the University of Calgary and the Canadian Centre for Computational Genomics.

The Care4Rare Canada Consortium has been funded since 2011 by Genome Canada and Ontario Genomics (OGI-049, OGI-064, OGI-147) in partnership with the Canadian Institutes of Health Research (CIHR ATE-110717, CIHR GPH-129346, CIHR GP1-155867, CIHR GPT-174518), Ontario Research Foundation, Genome Alberta, Genome British Columbia, Génome Québec, Children's Hospital of Eastern Ontario Foundation, Alberta Children's Hospital Research Institute, Alberta Innovates - Health Solutions, BC Children's Hospital Foundation, BC Children's Research Institute, BC Provincial Health Services Authority, Hospital for Sick Children, McLaughlin Centre for Molecular Medicine, and Phenotips Inc. K.M.B. is supported by a CIHR Foundation Grant (FDN-154279) and a Tier 1 Canada Research Chair in Rare Disease Precision Health. T.H. is supported by a CIHR Doctoral Award - Frederick Banting and Charles Best Canada Graduate Scholarship. H.H. is supported by the CHEO Foundation. See supplemental information for Consortium details.

Declarations of interests

M.B. has an equity interest in PhenoTips, which licenses software used in the Genomics4RD database. The remaining authors declare no competing interests.

Web resources

Canadian RDMM Network, http://rare-diseases-catalyst-network.ca/
Care4Rare Canada, http://www.care4rare.ca/
DECIPHER, https://www.deciphergenomics.org/
PhenomeCentral, www.phenomecentral.org/
GeneMatcher, https://genematcher.org/
Genome-wide Sequencing Ontario, https://gsontario.ca/
Genomics4RD, https://www.genomics4rd.ca/
Global Alliance for Genomics and Health, https://www.ga4gh.org/
International Rare Disease Research Consortium, https://irdirc.org/
Matchmaker Exchange, https://www.matchmakerexchange.org/
RD-Connect GPAP, https://rd-connect.eu/what-we-do/omics/gpap/
RDMM International, https://rdmminternational.org/
seqr, https://seqr.broadinstitute.org/
World Economic Forum, https://www.weforum.org/

References

1. Boycott, K.M., Vanstone, M.R., Bulman, D.E., and MacKenzie, A.E. (2013). Rare-disease genetics in the era of next-generation sequencing: discovery to translation. Nat. Rev. Genet. *14*, 681–691.
2. Beaulieu, C.L., Majewski, J., Schwartzentruber, J., Samuels, M.E., Fernandez, B.A., Bernier, F.P., Brudno, M., Knoppers, B., Marcadier, J., Dyment, D., et al. (2014). FORGE Canada Consortium: outcomes of a 2-year national rare-disease gene-discovery project. Am. J. Hum. Genet. *94*, 809–817.
3. Wright, C.F., McRae, J.F., Clayton, S., Gallone, G., Aitken, S., FitzGerald, T.W., Jones, P., Prigmore, E., Rajan, D., Lord, J., et al. (2018). Making new genetic diagnoses with old data: iterative reanalysis and reporting from genome-wide data in 1,133 families with developmental disorders. Genet. Med. *20*, 1216–1223.
4. Baxter, S.M., Posey, J.E., Lake, N.J., Sobreira, N., Chong, J.X., Buyske, S., Blue, E.E., Chadwick, L.H., Coban-Akdemir, Z.H., Doheny, K.F., et al. (2021). Centers for Mendelian Genomics: a decade of facilitating gene discovery. Genet. Med. *24*, 784–797.
5. Lochmüller, H., Badowska, D.M., Thompson, R., Knoers, N.V., Aartsma-Rus, A., Gut, I., Wood, L., Harmuth, T., Durudas, A., Graessner, H., et al. (2018). RD-Connect, NeurOmics and EURenOmics: collaborative European initiative for rare diseases. Eur. J. Hum. Genet. *26*, 778–785.
6. Majewski, J., Schwartzentruber, J.A., Caqueret, A., Patry, L., Marcadier, J., Fryns, J.P., Boycott, K.M., Ste-Marie, L.G., Mckiernan, F.E., Marik, I., et al. (2011). Mutations in *NOTCH2* in families with Hajdu-Cheney syndrome. Hum. Mutat. *32*, 1114–1117.
7. Bernier, F.P., Caluseriu, O., Ng, S., Schwartzentruber, J., Buckingham, K.J., Innes, A.M., Jabs, E.W., Innis, J.W., Schuette, J.L., Gorski, J.L., et al. (2012). Haploinsufficiency of *SF3B4*, a component of the pre-mRNA spliceosomal complex, causes Nager syndrome. Am. J. Hum. Genet. *90*, 925–933.
8. Hood, R.L., Lines, M.A., Nikkel, S.M., Schwartzentruber, J., Beaulieu, C., Nowaczyk, M.J.M., Allanson, J., Kim, C.A., Wieczorek, D., Moilanen, J.S., et al. (2012). Mutations in *SRCAP*, encoding SNF2-related CREBBP activator protein, cause Floating-Harbor syndrome. Am. J. Hum. Genet. *90*, 308–313.
9. Gibson, W.T., Hood, R.L., Zhan, S.H., Bulman, D.E., Fejes, A.P., Moore, R., Mungall, A.J., Eydoux, P., Babul-Hirji, R., An, J., et al. (2012). Mutations in *EZH2* cause Weaver syndrome. Am. J. Hum. Genet. *90*, 110–118.
10. Rivière, J.B., Mirzaa, G.M., O'Roak, B.J., Beddaoui, M., Alcantara, D., Conway, R.L., St-Onge, J., Schwartzentruber, J.A., Gripp, K.W., Nikkel, S.M., et al. (2012). De novo germline and postzygotic mutations in *AKT3, PIK3R2* and *PIK3CA* cause a spectrum of related megalencephaly syndromes. Nat. Genet. *44*, 934–940.
11. Doherty, D., Chudley, A.E., Coghlan, G., Ishak, G.E., Innes, A.M., Lemire, E.G., Rogers, R.C., Mhanni, A.A., Phelps, I.G., Jones, S.J.M., et al. (2012). *GPSM2* mutations cause the brain malformations and hearing loss in Chudley-McCullough syndrome. Am. J. Hum. Genet. *90*, 1088–1093.
12. Lines, M.A., Huang, L., Schwartzentruber, J., Douglas, S.L., Lynch, D.C., Beaulieu, C., Guion-Almeida, M.L., Zechi-Ceide, R.M., Gener, B., Gillessen-Kaesbach, G., et al. (2012). Haploinsufficiency of a spliceosomal GTPase encoded by *EFTUD2* causes mandibulofacial dysostosis with microcephaly. Am. J. Hum. Genet. *90*, 369–377.
13. Dyment, D.A., Smith, A.C., Alcantara, D., Schwartzentruber, J.A., Basel-Vanagaite, L., Curry, C.J., Temple, I.K., Reardon, W., Mansour, S., Haq, M.R., et al. (2013). Mutations in *PIK3R1* cause SHORT syndrome. Am. J. Hum. Genet. *93*, 158–166.
14. Srour, M., Schwartzentruber, J., Hamdan, F.F., Ospina, L.H., Patry, L., Labuda, D., Massicotte, C., Dobrzeniecka, S., Capo-Chichi, J.M., Papillon-Cavanagh, S., et al. (2012). Mutations in *C5ORF42* cause Joubert syndrome in the French Canadian population. Am. J. Hum. Genet. *90*, 693–700.
15. Srour, M., Hamdan, F.F., Schwartzentruber, J.A., Patry, L., Ospina, L.H., Shevell, M.I., Désilets, V., Dobrzeniecka, S., Mathonnet, G., Lemyre, E., et al. (2012). Mutations in *TMEM231* cause Joubert syndrome in French Canadians. J. Med. Genet. *49*, 636–641.
16. McDonell, L.M., Mirzaa, G.M., Alcantara, D., Schwartzentruber, J., Carter, M.T., Lee, L.J., Clericuzio, C.L., Graham, J.M., Morris-Rosendahl, D.J., Polster, T., et al. (2013). Mutations in *STAMBP*, encoding a deubiquitinating enzyme, cause microcephaly-capillary malformation syndrome. Nat. Genet. *45*, 556–562.
17. Lynch, D.C., Revil, T., Schwartzentruber, J., Bhoj, E.J., Innes, A.M., Lamont, R.E., Lemire, E.G., Chodirker, B.N., Taylor, J.P., Zackai, E.H., et al. (2014). Disrupted auto-regulation of the spliceosomal gene *SNRPB* causes cerebro-costo-mandibular syndrome. Nat. Commun. *5*, 4483.
18. Bennett, J.T., Tan, T.Y., Alcantara, D., Tétrault, M., Timms, A.E., Jensen, D., Collins, S., Nowaczyk, M.J.M., Lindhurst, M.J., Christensen, K.M., et al. (2016). Mosaic activating mutations in *FGFR1* cause encephalocraniocutaneous lipomatosis. Am. J. Hum. Genet. *98*, 579–587.
19. Srour, M., Hamdan, F.F., McKnight, D., Davis, E., Mandel, H., Schwartzentruber, J., Martin, B., Patry, L., Nassif, C., Dionne-Laporte, A., et al. (2015). Joubert syndrome in French Canadians and identification of mutations in *CEP104*. Am. J. Hum. Genet. *97*, 744–753.
20. Nikkel, S.M., Dauber, A., de Munnik, S., Connolly, M., Hood, R.L., Caluseriu, O., Hurst, J., Kini, U., Nowaczyk, M.J.M., Afenjar, A., et al. (2013). The phenotype of Floating-Harbor syndrome: clinical characterization of 52 individuals with mutations in exon 34 of *SRCAP*. Orphanet J. Rare Dis. *8*, 63.
21. Huang, L., Vanstone, M.R., Hartley, T., Osmond, M., Barrowman, N., Allanson, J., Baker, L., Dabir, T.A., Dipple, K.M., Dobyns, W.B., et al. (2016). Mandibulofacial dysostosis with

microcephaly: mutation and database update. Hum. Mutat. *37*, 148–154.
22. Hood, R.L., Schenkel, L.C., Nikkel, S.M., Ainsworth, P.J., Pare, G., Boycott, K.M., Bulman, D.E., and Sadikovic, B. (2016). The defining DNA methylation signature of Floating-Harbor syndrome. Sci. Rep. *6*, 38803.
23. Balci, T.B., Hartley, T., Xi, Y., Dyment, D.A., Beaulieu, C.L., Bernier, F.P., Dupuis, L., Horvath, G.A., Mendoza-Londono, R., Prasad, C., et al. (2017). Debunking Occam's razor: diagnosing multiple genetic diseases in families by whole-exome sequencing. Clin. Genet. *92*, 281–289.
24. Sawyer, S.L., Hartley, T., Dyment, D.A., Beaulieu, C.L., Schwartzentruber, J., Smith, A., Bedford, H.M., Bernard, G., Bernier, F.P., Brais, B., et al. (2016). Utility of whole-exome sequencing for those near the end of the diagnostic odyssey: time to address gaps in care. Clin. Genet. *89*, 275–284.
25. Boycott, K.M., Hartley, T., Biesecker, L.G., Gibbs, R.A., Innes, A.M., Riess, O., Belmont, J., Dunwoodie, S.L., Jojic, N., Lassmann, T., et al. (2019). A diagnosis for all rare genetic diseases: the horizon and the next frontiers. Cell *177*, 32–37.
26. Hartley, T., Lemire, G., Kernohan, K.D., Howley, H.E., Adams, D.R., and Boycott, K.M. (2020). New diagnostic approaches for undiagnosed rare genetic diseases. Annu. Rev. Genom. Hum. Genet. *21*, 351–372.
27. Dyment, D.A., O'Donnell-Luria, A., Agrawal, P.B., Coban Akdemir, Z., Aleck, K.A., Antaki, D., Al Sharhan, H., Au, P.Y.B., Aydin, H., Beggs, A.H., et al. (2021). Alternative genomic diagnoses for individuals with a clinical diagnosis of Dubowitz syndrome. Am. J. Med. Genet. *185*, 119–133.
28. Jones, K.L., Jones, M.C., and del Campo, M. (2021). Smith's Recognizable Patterns of Human Malformation (Elsevier Saunders).
29. Girdea, M., Dumitriu, S., Fiume, M., Bowdin, S., Boycott, K.M., Chénier, S., Chitayat, D., Faghfoury, H., Meyn, M.S., Ray, P.N., et al. (2013). PhenoTips: patient phenotyping software for clinical and research use. Hum. Mutat. *34*, 1057–1065.
30. Buske, O.J., Girdea, M., Dumitriu, S., Gallinger, B., Hartley, T., Trang, H., Misyura, A., Friedman, T., Beaulieu, C., Bone, W.P., et al. (2015). PhenomeCentral: a portal for phenotypic and genotypic matchmaking of patients with rare genetic diseases. Hum. Mutat. *36*, 931–940.
31. Osmond, M., Hartley, T., Johnstone, B., Andjic, S., Girdea, M., Gillespie, M., Buske, O., Dumitriu, S., Koltunova, V., Ramani, A., et al. (2022). PhenomeCentral: 7 years of rare disease matchmaking. Hum. Mutat. *43*, 674–681.
32. Philippakis, A.A., Azzariti, D.R., Beltran, S., Brookes, A.J., Brownstein, C.A., Brudno, M., Brunner, H.G., Buske, O.J., Carey, K., Doll, C., et al. (2015). The Matchmaker Exchange: a platform for rare disease gene discovery. Hum. Mutat. *36*, 915–921.
33. Hamosh, A., Wohler, E., Martin, R., Griffith, S., Rodrigues, E.d.S., Antonescu, C., Doheny, K.F., Valle, D., and Sobreira, N. (2022). The impact of GeneMatcher on international data sharing and collaboration. Hum. Mutat. *43*, 668–673.
34. Foreman, J., Brent, S., Perrett, D., Bevan, A.P., Hunt, S.E., Cunningham, F., Hurles, M.E., and Firth, H.V. (2022). DECIPHER: supporting the interpretation and sharing of rare disease phenotype-linked variant data to advance diagnosis and research. Hum. Mutat. *43*, 682–697.
35. Laurie, S., Piscia, D., Matalonga, L., Corvó, A., Fernández-Callejo, M., Garcia-Linares, C., Hernandez-Ferrer, C., Luengo, C., Martínez, I., Papakonstantinou, A., et al. (2022). The RD-Connect Genome-Phenome Analysis Platform: accelerating diagnosis, research, and gene discovery for rare diseases. Hum. Mutat. *43*, 717–733.
36. Pais, L.S., Snow, H., Weisburd, B., Zhang, S., Baxter, S.M., DiTroia, S., O'Heir, E., England, E., Chao, K.R., Lemire, G., et al. (2021). seqr : a web-based analysis and collaboration tool for rare disease genomics. Hum. Mutat. *43*, 698–707.
37. Boycott, K.M., Azzariti, D.R., Hamosh, A., and Rehm, H.L. (2022). Seven years since the launch of the Matchmaker Exchange: the evolution of genomic matchmaking. Hum. Mutat. *43*, 659–667.
38. Faden, M., Alzahrani, F., Mendoza-Londono, R., Dupuis, L., Hartley, T., Kannu, P., Raiman, J.A., Howard, A., Qin, W., Tetreault, M., et al. (2015). Identification of a recognizable progressive skeletal dysplasia caused by *RSPRY1* mutations. Am. J. Hum. Genet. *97*, 608–615.
39. Simons, C., Dyment, D., Bent, S.J., Crawford, J., D'Hooghe, M., Kohlschütter, A., Venkateswaran, S., Helman, G., Poll-The, B.T., Makowski, C.C., et al. (2017). A recurrent de novo mutation in *TMEM106B* causes hypomyelinating leukodystrophy. Brain *140*, 3105–3111.
40. Kernohan, K.D., Dyment, D.A., Pupavac, M., Cramer, Z., McBride, A., Bernard, G., Straub, I., Tetreault, M., Hartley, T., Huang, L., et al. (2017). Matchmaking facilitates the diagnosis of an autosomal-recessive mitochondrial disease caused by biallelic mutation of the tRNA isopentenyltransferase (*TRIT1*) gene. Hum. Mutat. *38*, 511–516.
41. Lee, Y.R., Khan, K., Armfield-Uhas, K., Srikanth, S., Thompson, N.A., Pardo, M., Yu, L., Norris, J.W., Peng, Y., Gripp, K.W., et al. (2020). Mutations in *FAM50A* suggest that Armfield XLID syndrome is a spliceosomopathy. Nat. Commun. *11*, 3698.
42. Vavassori, S., Chou, J., Faletti, L.E., Haunerdinger, V., Opitz, L., Joset, P., Fraser, C.J., Prader, S., Gao, X., Schuch, L.A., et al. (2021). Multisystem inflammation and susceptibility to viral infections in human ZNFX1 deficiency. J. Allergy Clin. Immunol. *148*, 381–393.
43. Salpietro, V., Dixon, C.L., Guo, H., Bello, O.D., Vandrovcova, J., Efthymiou, S., Maroofian, R., Heimer, G., Burglen, L., Valence, S., et al. (2019). AMPA receptor GluA2 subunit defects are a cause of neurodevelopmental disorders. Nat. Commun. *10*, 3094.
44. Lessel, D., Zeitler, D.M., Reijnders, M.R.F., Kazantsev, A., Hassani Nia, F., Bartholomäus, A., Martens, V., Bruckmann, A., Graus, V., McConkie-Rosell, A., et al. (2020). Germline *AGO2* mutations impair RNA interference and human neurological development. Nat. Commun. *11*, 5797.
45. Chelban, V., Wilson, M.P., Warman Chardon, J., Vandrovcova, J., Zanetti, M.N., Zamba-Papanicolaou, E., Efthymiou, S., Pope, S., Conte, M.R., Abis, G., et al. (2019). *PDXK* mutations cause polyneuropathy responsive to pyridoxal 5'-phosphate supplementation. Ann. Neurol. *86*, 225–240.
46. den Hoed, J., de Boer, E., Voisin, N., Dingemans, A.J.M., Guex, N., Wiel, L., Nellaker, C., Amudhavalli, S.M., Banka, S., Bena, F.S., et al. (2021). Mutation-specific pathophysiological mechanisms define different neurodevelopmental disorders associated with SATB1 dysfunction. Am. J. Hum. Genet. *108*, 346–356.
47. Au, P.Y.B., You, J., Caluseriu, O., Schwartzentruber, J., Majewski, J., Bernier, F.P., Ferguson, M., Care for Rare Canada Consortium, Valle, D., Parboosingh, J.S., et al. (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*. Hum. Mutat. *36*, 1009–1014.
48. Osmond, M., Hartley, T., Dyment, D.A., Kernohan, K.D., Brudno, M., Buske, O.J., Innes, A.M., Boycott, K.M., Care4Rare Canada Consortium, and Bernier, F., et al. (2022). Outcome of over 1500 matches through the Matchmaker Exchange for rare disease gene discovery: the 2-year experience of Care4Rare Canada. Genet. Med. *24*, 100–108.
49. Boycott, K.M., Campeau, P.M., Howley, H.E., Pavlidis, P., Rogic, S., Oriel, C., Berman, J.N., Hamilton, R.M., Hicks, G.G., Lipshitz, H.D., et al. (2020). The Canadian Rare Diseases Models and Mechanisms (RDMM) network: connecting understudied genes to model organisms. Am. J. Hum. Genet. *106*, 143–152.
50. Frosk, P., Arts, H.H., Philippe, J., Gunn, C.S., Brown, E.L., Chodirker, B., Simard, L., Majewski, J., Fahiminiya, S., Russell, C., et al. (2017). A truncating mutation in *CEP55* is the likely cause of MARCH, a novel syndrome affecting neuronal mitosis. J. Med. Genet. *54*, 490–501.
51. Barrie, E.S., Overwater, E., van Haelst, M.M., Motazacker, M.M., Truxal, K.V., Crist, E., Mostafavi, R., Pivnick, E.K., Choudhri, A.F., Narumanchi, T., et al. (2020). Expanding the spectrum of *CEP55*-associated disease to viable phenotypes. Am. J. Med. Genet. *182*, 1201–1208.
52. Driver, H.G., Hartley, T., Price, E.M., Turinsky, A.L., Buske, O.J., Osmond, M., Ramani, A.K., Kirby, E., Kernohan, K.D., Couse, M., et al. (2022). Genomics4RD: an integrated platform to share Canadian deep-phenotype and multi-omic data for international rare disease gene discovery. Hum. Mutat. *43*, 800–811.
53. Woolley, J.P., Kirby, E., Leslie, J., Jeanson, F., Cabili, M.N., Rushton, G., Hazard, J.G., Ladas, V., Veal, C.D., Gibson, S.J., et al. (2018). Responsible sharing of biomedical data and biospecimens via the "Automatable Discovery and Access Matrix" (ADA-M). NPJ Genom. Med. *3*, 17.
54. Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W.W., Hegde, M., Lyon, E., Spector, E., et al. (2015). Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet. Med. *17*, 405–424.
55. Strande, N.T., Riggs, E.R., Buchanan, A.H., Ceyhan-Birsoy, O., DiStefano, M., Dwight, S.S., Goldstein, J., Ghosh, R., Seifert, B.A., Sneddon, T.P., et al. (2017). Evaluating the clinical validity of gene-disease associations: an evidence-based framework developed by the Clinical Genome Resource. Am. J. Hum. Genet. *100*, 895–906.
56. Landrum, M.J., Chitipiralla, S., Brown, G.R., Chen, C., Gu, B., Hart, J., Hoffman, D., Jang, W., Kaur, K., Liu, C., et al. (2020). ClinVar: improvements to accessing data. Nucleic Acids Res. *48*, D835–D844.
57. Fokkema, I.F.A.C., Taschner, P.E.M., Schaafsma, G.C.P., Celli, J., Laros, J.F.J., and den Dunnen, J.T. (2011). LOVD v.2.0: the next generation in gene variant databases. Hum. Mutat. *32*, 557–563.
58. Zawati, M.H., Parry, D., Thorogood, A., Nguyen, M.T., Boycott, K.M., Rosenblatt, D., and Knoppers, B.M. (2014). Reporting results from whole-genome and whole-exome sequencing in clinical practice: a proposal for Canada? J. Med. Genet. *51*, 68–70.
59. Kleiderman, E., Knoppers, B.M., Fernandez, C.V., Boycott, K.M., Ouellette, G., Wong-Rieger, D., Adam, S., Richer, J., and Avard, D. (2014). Returning incidental findings from genetic research to children: views of parents of children affected by rare diseases. J. Med. Ethics *40*, 691–696.
60. Boycott, K., Hartley, T., Adam, S., Bernier, F., Chong, K., Fernandez, B.A., Friedman, J.M., Geraghty, M.T., Hume, S., Knoppers, B.M., et al. (2015). The clinical application of genome-wide sequencing for monogenic diseases in Canada: position statement of the Canadian College of Medical Geneticists. J. Med. Genet. *52*, 431–437.
61. Esquivel-Sada, D., and Nguyen, M.T. (2018). Diagnosis of rare diseases under focus: impacts for Canadian patients. J. Community Genet. *9*, 37–50.
62. Marshall, D.A., MacDonald, K.V., Heidenreich, S., Hartley, T., Bernier, F.P., Gillespie, M.K., McInnes, B., Innes, A.M., Armour, C.M., and Boycott, K.M. (2019). The value of diagnostic testing for parents of children with rare genetic diseases. Genet. Med. *21*, 2798–2806.
63. Hayeems, R.Z., Marshall, C.R., Gillespie, M.K., Szuto, A., Chisholm, C., Stavropoulos, D.J., Venkataramanan, V., Tsiplova, K., Sawyer, S., Price, E.M., et al. (2022). Comparing genome sequencing technologies to improve rare disease diagnostics: a protocol for the evaluation of a pilot project, Genome-wide Sequencing Ontario. CMAJ Open *10*, E460–E465.
64. Lazier, J., Hartley, T., Brock, J.-A., Caluseriu, O., Chitayat, D., Laberge, A.-M., Langlois, S., Lauzon, J., Nelson, T.N., Parboosingh, J., et al. (2021). Clinical application of fetal genome-wide sequencing during pregnancy: position statement of the Canadian College of Medical Geneticists. J. Med. Genet. *59*, 931–937.
65. Eaton, A., Hartley, T., Kernohan, K., Ito, Y., Lamont, R., Parboosingh, J., Barrowman, N., Care4Rare consortium, Innes, A.M., and Boycott, K. (2020). When to think outside the autozygome: best practices for exome sequencing in "consanguineous" families. Clin. Genet. *97*, 835–843.
66. Hartley, J.N., Simard, L.R., Ly, V., Del Bigio, M.R., and Frosk, P. (2019). A homozygous canonical splice acceptor site mutation in *PRUNE1* is responsible for a rare childhood neurodegenerative disease in Manitoba Cree families. Am. J. Med. Genet. *179*, 206–218.
67. Choquet, K., Tétreault, M., Yang, S., La Piana, R., Dicaire, M.J., Vanstone, M.R., Mathieu, J., Bouchard, J.P., Rioux, M.F., Rouleau, G.A., et al. (2016). *SPG7* mutations explain a significant proportion of French Canadian spastic ataxia cases. Eur. J. Hum. Genet. *24*, 1016–1021.
68. Lovric, S., Goncalves, S., Gee, H.Y., Oskouian, B., Srinivas, H., Choi, W.I., Shril, S., Ashraf, S., Tan, W., Rao, J., et al. (2017). Mutations in sphingosine-1-phosphate lyase cause nephrosis with ichthyosis and adrenal insufficiency. J. Clin. Invest. *127*, 912–928.
69. Boycott, K.M., Beaulieu, C.L., Kernohan, K.D., Gebril, O.H., Mhanni, A., Chudley, A.E., Redl, D., Qin, W., Hampson, S., Küry, S., et al. (2015). Autosomal-recessive intellectual disability with cerebellar atrophy syndrome caused by mutation of the manganese and zinc transporter gene *SLC39A8*. Am. J. Hum. Genet. *97*, 886–893.
70. Canadian Organization for Rare Disorders (CORD) (2015). Now is the time: a strategy for rare disease is a strategy for all Canadians. https://www.raredisorders.ca/content/uploads/CORD_Canada_RD_Strategy_22May15.pdf.
71. Marshall, D.A., Benchimol, E.I., MacKenzie, A., Duque, D.R., MacDonald, K.V., Hartley, T., Howley, H., Hamilton, A.,

Gillespie, M., Malam, F., and Boycott, K.M. (2019). Direct health-care costs for children diagnosed with genetic diseases are significantly higher than for children with other chronic diseases. Genet. Med. *21*, 1049–1057.

72. Rehm, H.L., Page, A.J.H., Smith, L., Adams, J.B., Alterovitz, G., Babb, L.J., Barkley, M.P., Baudis, M., Beauvais, M.J.S., Beck, T., et al. (2021). GA4GH: international policies and standards for data sharing across genomic research and healthcare. Cell Genom. *1*, 100029.
73. Boycott, K.M., Rath, A., Chong, J.X., Hartley, T., Alkuraya, F.S., Baynam, G., Brookes, A.J., Brudno, M., Carracedo, A., den Dunnen, J.T., et al. (2017). International cooperation to enable the diagnosis of all rare genetic diseases. Am. J. Hum. Genet. *100*, 695–705.
74. Chediak, L. (2020). Sharing Sensitive Health Data in a Federated Data Consortium Model: An Eight-step Guide (World Economic Forum). https://www3.weforum.org/docs/WEF_Sharing_Sensitive_Health_Data_2020.pdf.
75. Boycott, K.M., Dyment, D.A., and Innes, A.M. (2018). Unsolved recognizable patterns of human malformation: Challenges and opportunities. Am. J. Med. Genet. C Semin. Med. Genet. *178*, 382–386.

The American Journal of Human Genetics, Volume *109*

Supplemental information

Care4Rare Canada: Outcomes from a decade of network science for rare disease gene discovery

Kym M. Boycott, Taila Hartley, Kristin D. Kernohan, David A. Dyment, Heather Howley, A. Micheil Innes, Francois P. Bernier, Michael Brudno, and Care4Rare Canada Consortium

Supplemental Information

Content

Box S1. Matchmaking ends a 19-year diagnostic odyssey by disentangling two novel rare diseases

The importance of global matchmaking for RD gene discovery, and ultimately diagnostic clarity for all affected families, cannot be overstated. In one striking example, we identified, in parallel, candidate variants in novel genes in a single individual. A baby boy presented with marked hypotonia and nystagmus shortly after birth. Brain MRI showed lack of myelination, and a likely clinical diagnosis of Pelizaeus-Merzbacher-like disease (MIM 312080) was given although *PLP1* (MIM 300401) testing was negative. A gastrostomy tube was placed at 5 months of age for a failure to thrive. Around that time, he developed intermittent episodes of choreoathetosis, initially thought to be seizures, evolving to dystonic arm posturing in his teen years. He has severe intellectual disability. Following research exome sequencing in 2015 as part of the Care4Rare project, we identified *de novo* variants in two genes that had never been associated with human disease, *TMEM106B* [exon 8; c.754G>A p.(Asp252Asn); NM_001134232.2] (MIM 613413) and *USP7* [exon 14; c.1406T>G p.(Val469Gly); NM_001286457.2] (MIM 602519). Both genes were entered into the Matchmaker Exchange data sharing platform. We matched immediately for *USP7* to an international cohort of 16 individuals with autism, intellectual disability, epilepsy, and feeding difficulties (MIM 616863).[1] No affected individuals, however, exhibited hypomyelination and so we kept our submission for *TMEM106B* active within MME. One year later, we matched to another proband with hypomyelination and the same *de novo* mutation in *TMEM106B*. Two years following our initial review of the exome data, we published a cohort of four affected individuals from Canada, Australia, Germany, and the Netherlands with the same recurrent *TMEM106B* mutation (MIM 617964).[2] Global matchmaking made it possible to disentangle the co-occurrence of two different RDs for this individual, while discovering two new RDs and characterizing their clinical spectrum, so that other families will not have to wait 19 years for diagnostic clarity. The family is now an active member of the USP7 Foundation, a non-profit organization dedicated to supporting families and research for this RD.

Box S2. Querying databases for variants in candidate genes facilitates RD gene discovery

Two French Canadian sibships from the Care4Rare project had a strikingly similar and complex clinical presentation characterized by hereditary spastic paraplegia (HSP), intellectual disability, and cerebral anomalies yet remained without a diagnosis following research exome sequencing in 2015. Re-analysis in 2020 with the latest Care4Rare bioinformatics pipeline revealed a shared homozygous candidate variant in *ABHD16A* (MIM 142620) that was missed by our earlier bioinformatics pipeline. We immediately submitted the gene to MME but no matches were returned. In 2021, we queried the data in Genomics4RD, and quickly identified two probands with rare, deleterious looking biallelic variants in *ABHD16A* and symptoms of complex HSP: one family who remained without a diagnosis and whose data had not yet been re-analyzed, and a second family whose data had just been returned from a US-based clinical laboratory. A subsequent query of rare variants in Geno2MP quickly identified two additional families from the USA. These findings add *ABHD16A* to the growing list of genes where deleterious variants cause early-onset autosomal recessive complicated HSP through the dysregulation of lipid metabolism (MIM 619735).[3] This discovery highlights the power of querying databases for variants in candidate genes in genomic data from families with unsolved RD.

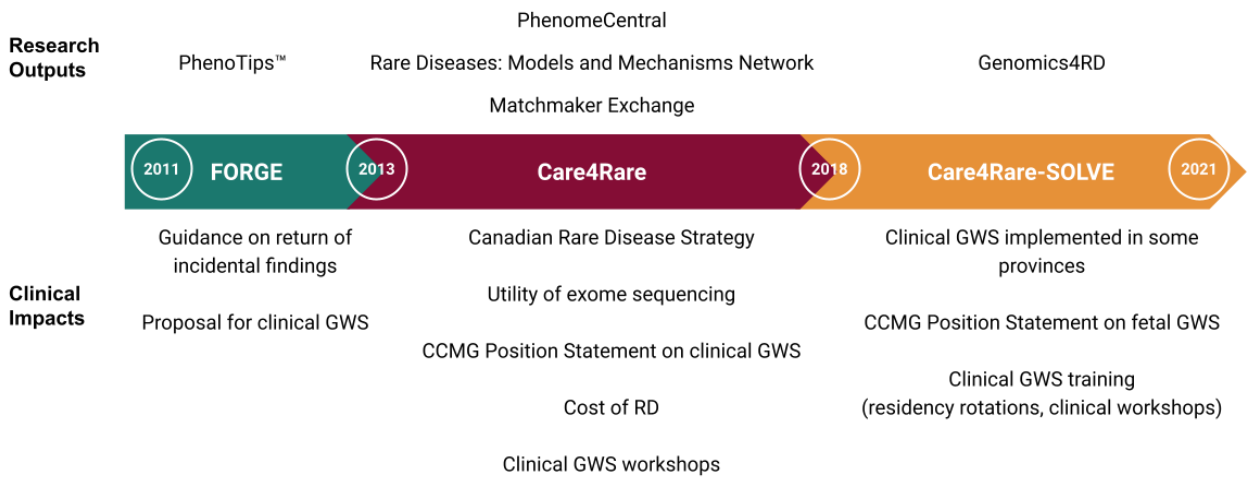

Figure S1. Research outputs and clinical impacts resulting from a decade of the Care4Rare Canada Consortium. In addition to its considerable efforts in rare disease (RD) diagnosis (n=623) and gene discovery (n=203), the Care4Rare Canada Consortium identified and responded to the needs of the RD research and clinical communities. Through all eras, the Consortium delivered key technologies, evidence, resources, and guides to support RD diagnosis and discovery. For the research community, the Care4Rare Consortium delivered technologies to support RD data capture (PhenoTips™) and sharing (PhenomeCentral, Matchmaker Exchange, Genomics4RD), as well as facilitated functional characterization research of novel genes in model organisms by establishing the Canadian Rare Diseases: Models and Mechanisms Network. For the clinical community, the Consortium (in partnership with key stakeholders) paved the way for clinical implementation of genome-wide sequencing (GWS) by publishing evidence of its diagnostic utility and the costs of RD; contributing to a national RD strategy, developing clinical guidelines and position statements; and building capacity through curriculum development and workshops. Ten years later, Canada remains an international leader in RD discovery and GWS is available as a publicly funded test for complex RD diagnosis in most regions of Canada.

Eras	FORGE	Care4Rare	SOLVE
General inclusion criteria	• Assessment by a clinical geneticist or subspecialist practicing in genetics who is a member of Care4Rare Canada Consortium • Suspected monogenic disease • Standard-of-care testing in their province/territory completed • No molecular diagnosis or compelling VUS • Consented using the Care4Rare Research Ethics Board-approved protocol • Samples available and follow-up possible • Family members available, including deep-phenotype and samples		
Population	• Children • At least one affected individual from Canada available for a particular project	Children and adults	Children and adults
Era specific criteria	• Highly recognizable syndrome • Family history of recurrence • Known consanguinity • Isolated community	FORGE criteria plus single families	Genomic data only AND/OR Multi-omics data

Table S1. Inclusion Criteria for Participation in Care4Rare across the Eras

Group	RD Characteristics	Reason(s) for Remaining Unsolved	Technologies Proposed
1	Recognizable RD with no or limited genetic heterogeneity (e.g., Tuberous sclerosis)	Pathogenic variation is undetectable with current technologies	Genome sequencing RNA sequencing +/-Deep sequencing
2	Recognizable RD where the molecular etiology has not yet been elucidated (e.g. PHACES, as reviewed by Boycott et al[4])	Molecular etiology is not yet known and is undetectable using current approaches OR Non-Mendelian inheritance OR Non-genetic	Large-scale data sharing Multi-omics Novel computational approaches
3	RD with high degree of genetic heterogeneity (e.g., cerebellar ataxia)	Pathogenic variant in a known gene that is missed by current diagnostic technologies or clinical panel OR Pathogenic variant in a novel gene	Re-analysis Genome sequencing RNA sequencing
4	RD associated with a constellation of features without a name	Phenotypic expansion of a known disease OR Novel gene-disease association OR More than one genetic diagnosis OR Non-Mendelian inheritance OR Non-genetic	Re-analysis Multi-omics Large-scale data sharing

Table S3. Overview of a Research Protocol for RDs Remaining Unsolved Following Exome Sequencing
Multi-omics refers to genome and RNA sequencing data, lipidomics, metabolomics, and/or epigenomics in any combination

Care4Rare Canada Consortium: Current and past leadership teams

FORGE Canada project: Finding of Rare Disease Genes in Canada

Kym Boycott (lead; University of Ottawa), Jan Friedman (co-lead; University of British Columbia), Jacques Michaud (co-lead; University of Montreal), Denise Avard (McGill University), Francois Bernier (University of Calgary), Michael Brudno (University of Toronto), Dennis Bulman (University of Ottawa), Bridget Fernandez (Memorial University of Newfoundland), Steven Jones (BC Cancer Agency Genome Sciences Centre), Bartha Knoppers (McGill University), Jacek Majewski (McGill University), Janet Marcadier (Children's Hospital of Eastern Ontario Research Institute), Marco Marra (BC Cancer Agency Genome Sciences Centre), Alexandre Montpetit (McGill Applied Genomics Innovation Core), Andrew Orr (Dalhousie University), Francois Rousseau (University of Laval), Mark Samuels (University of Montreal), Steven Scherer (The Centre for Applied Genomics), Richard Wintle (The Centre for Applied Genomics).

Care4Rare project: Enhanced Care for Rare Genetic Diseases in Canada

Kym Boycott (lead; University of Ottawa), Alex MacKenzie (co-lead; University of Ottawa), Genevieve Bernard (McGill University), Francois Bernier (University of Calgary), Bernard Brais (McGill University), Michael Brudno (University of Toronto), Dennis Bulman (University of Ottawa), David Dyment (University of Ottawa), Roy Gravel (University of Calgary), Micheil Innes (University of Calgary), Bartha Knoppers (McGill University), Jacek Majewski (McGill University), Deborah Marshall (University of Calgary), Christopher McMaster (Dalhousie University), Jacques Michaud (University of Montreal), Mark Samuels (University of Montreal), Barbara von Tigerstrom (University of Saskatchewan).

Care4Rare SOLVE project: Care4Rare Canada: Harnessing Multi-omics to Deliver Innovative Diagnostic Care for Rare Genetic Diseases in Canada

Kym Boycott (lead; University of Ottawa), Michael Brudno (co-lead, University of Toronto), Francois Bernier (co-lead, University of Calgary), Clara van Karnebeek (co-lead, University of British Columbia), James Dowling (University of Toronto), David Dyment (CHEO Research Institute), Robin Hayeems (University of Toronto), Micheil Innes (University of Calgary), Kristin Kernohan (University of Ottawa), Bartha Knoppers (McGill University), Ryan Lamont (University of Calgary), Anna Lehman (University of British Columbia), Christian Marshall (University of Toronto), Deborah Marshall (University of Calgary), Roberto Mendoza (University of Toronto), Sara Mostafavi (University of British Columbia), Jillian Parboosingh (University of Calgary).

Care4Rare Canada Administrative Team

Project management: Taila Hartley (Operations Director, University of Ottawa; current), Elisabeth Soubry (Project Manager, University of Ottawa; current), Meryl Akers (Project Manager, University of Toronto; current), Brenda McInnes (Project Manager, Univesity of Calgary; current), Meriam Waqas (Project Manager, University of British Columbia; current), Hanh Dao (Administrative Assistant, University of Ottawa; current), Grace Ediae (Project Manager, University of Ottawa; past), Chandree Beaulieu (Project Manager, University of Ottawa; past), Janet Marcadier (Project Manager, University of Ottawa; past).

Research Administration: Heather Howley (Research Funding Development Officer, current), Laura Minaker (Senior Finance Officer, current), Samira Chamaa (Manager of Grants and Awards, current).

Members of the Care4Rare Canada Consortium

Iman Abumansour, Nassima Addour, Sohnee Ahmed, Ebba Alkhunaizi, Judith Allanson, Kim Amburgey, Ingrid Ambus, Laura Arbour, Christine Armour, Linlea Armstrong, Taryn Athey, Setareh Ashtiani, Billie Au, Gudren Aubertin, Rita Aul, Lauren Badalato, Steven Baker, Tugce Balci, Torben Bech-Hansen, Karen Bedard, Melanie Bedford, Steffany Bennett, Samuel Berkovic, Genevieve Bernard, Priya Bhola, Neal Boerkoel, Brittney Bosse, Danielle Bourque, Sarah Bowdin, Lauren Brady, Tamara Braid, Bernard Brais, Paula Brna, Lina Buelvas, Dennis Bulman, Oana Caluseriu, Craig Campbell, Maggie Campbell, Philippe Campeau, Melissa Carter, Michelle Caudle, Inara Chacon, Lauren Chad, Pranesh Chakraborty, Erin Chamberlain, Alicia Chan, Caitlin Chang, Marisa Chard, Stephanie Chieffo, Caitlin Chisholm, David Chitayat, Tillie Chiu, Bernie Chodirker, Karen Chong, Albert Chudley, Mireille Cloutier, Ronni Cohn, Samantha Coliacovo, Ashley Collier, Melissa Cornthwaite, Patrick Cossette, Gregory Costain, Gina Cowing, Alessandra Cumming, Lauren Currie, Donna Cushing, Daniela D'Agostino, Dawn Danielson, Isabelle De Bie, Nomazulu Dlamini, Asif Doja, James Dowling, Saunya Dover, Eric Dupas, Lucie Dupuis, Sarah Dyack, David Dyment, Alison Eaton, Hanna Faghfoury, Sandra Farrell, Bridget Fernandez, Patrick Ferreira, Rachel Ferrier, Isabel Filges, Cynthia Forster-Gibson, Zoey Freedman, Jan Friedman, Patrick Frosk, Charlotte Fung, Michael Geraghty, Brenda Gerull, William Gibson, Meredith Gillespie, Jane Gillis, Cathy Gilpin, Nicole Gojska, Elaine Goh, Claire Goldsmith, Hernan Gonorazky, Sharan Goobie, Robert Gow, Sara Gracie, Gail Graham, Jane Green, Cheryl Greenberg, Eyal Grunebaum, Andrea Guerin, Elie Haddad, Farah Hassan, Elise Heon, Stacy Hewson, Christina Honeywell, Gabriella Horvath, Rebecca

Hough, Cathy Huculak, Stacey Hume, Peter Humphreys, Teresa Hunt, Alasdair Hunter, Stephanie Huynh, Michal Inbar-Feigenberg, Angela Inglis, Micheil Innes, Nada Jabado, Shailly Jain, Leslie Colvin James, Rebekah Jobling, Jack Jung, Roman Jurencak, Rita Kadida, Peter Kannu, Natalya Karp, Kathryn Keely, Kristin Kernohan, Aneal Khan, Raymond Kim, Courtney Kiss, Robert Koenekoop, Nataly Laflamme, Sylvie Langlois, Julie Lauzon, Jessica Lawrence, Joanna Lazier, Anna Lehman, Ordan Lehmann, Edmond Lemire, Gabrielle Lemire, Chumei Li, Matt Lines, Brian Lowry, Ian MacDonald, Karen MacDonald, Sara MacKay, Alex Mackenzie, Yolanda MacKinnon, Linda MacLaren, Joanna MacLean, Melissa MacPherson, Bruno Maranda, Julien Marcadier, Sandra Marles, Nicole Martin, Andre Mattman, Anna Matveev, Ashish Marwaha, Barbara McGillivray, Brenda McInnes, Margaret McKinnon, Ross McLeod, Christopher McMaster, Hugh McMillan, Kirsten Meagher, Jariwala Mehul, Roberto Mendoza, Saadet Mercimek-Andrews, Leanne Mercer, Wendy Meschino, Aziz Mhanni, Jacques Michaud, Kamiar Mireskandari, Burcin Morali, Chantal Morel, Stephen Mosca, Alexio Muise, Jillian Murphy, Mitzi Murray, Melanie Napier, Marjan Nezarati, Karen Niederhoffer, Sarah Nikkel, Graeme Nimmo, Malgorzata Nowaczyk, Laura Nunez, Margaret O'Riley, Andrew Orr, Matthew Osmond, Sandhya Parkash, Millan Patel, Melanie Paterson, Larissa Peck, Lynette Penny, Helene Perras, Renee Perrier, Anne Pham-huy, Chitra Prasad, Magda Price, Josee Provost, Julian Raiman, Julie Richer, Christie Riddell, Carleigh Robertson, Marie-Eve Robinson, Johane Robitaille, Maian Roifman, Samantha Rojas, David Rosenblatt, Alison Rusnak, Maha Saleh, Mark Samuels, Parkash Sandhya, Sarah Sawyer, Andreas Schulze, Kim Seath, Sheiva Shaari, Priyana Sharma, Eric Shoubridge, Komudi Siriwardena, Sandra Sirrs, Victoria Siu, David Skidmore, Carter Snead, Renata Sobiesiak, Neal Sondheimer, Taylor Speziale, Rebecca Sparkes, Myriam Srour, Sylvia Stockler, Julia Su, Oksana Suchowersky, Anna Szuto, Marta Szybowska, Julia Tagoe, Mark Tarnopolsky, Deborah Terespolsky, Maryann Thomas, Alan Tinmouth, Eva Tomiak, Clara van Karnebeek, Anthony Vandersteen, Kalene Van Engelen, Lea Velsher, Sunita Venkateswaran, Krista Vincent, Jagdeep Walia, Karin Wallace, Jodi Warman, Jonathan Wasserman, Rosanna Weksberg, Grace Yoon, Dana Young, Andrea Yu, Farah Zahir, Jessica Zambonin

Supplemental Acknowledgements

While the success of the Care4Rare Canada program is due innumerable individual efforts, there are key individuals who warrant special recognition: Dennis Bulman (Alberta Precision Laboratories) and Alex MacKenzie (CHEO Research Institute) for providing key early mentorship and support to launch the FORGE program; Cindy Bell and Marc LePage (Genome Canada) and Paul Lasko (CIHR) who were instrumental in developing funding programs to leverage genomics for RD discovery and diagnosis;

Martin Osmond (CHEO Research Institute) for ensuring the host institution was robustly able to support a national network for a decade; and Kevin Keohane (CHEO Foundation) for his championship of RD as an important area to study. We would also like to thank all the former and current trainees of the Consortium for their curiosity and dedication to RD research and the Canadian Organization for Rare Disorders for their valued partnership and support over the past decade.

Web Resources

Justin's Odyssey: A journey in rare disease data sharing - YouTube, https://www.youtube.com/watch?v=IDEhdzvlR4I)

Supplemental References

1. Fountain, M. D., Oleson, D. S., Rech, M. E., Segebrecht, L., Hunter, J. V., McCarthy, J. M., Lupo, P. J., Holtgrewe, M., Moran, R., Rosenfeld, J. A., et al. (2019). Pathogenic variants in *USP7* cause a neurodevelopmental disorder with speech delays, altered behavior, and neurologic anomalies. Genet. Med. *21*, 1797–1807.
2. Simons, C., Dyment, D., Bent, S.J., Crawford, J., D'Hooghe, M., Kohlschütter, A., Venkateswaran, S., Helman, G., Poll-The, B.T., Makowski, C.C., et al. (2017). A recurrent de novo mutation in *TMEM106B* causes hypomyelinating leukodystrophy. Brain *140*, 3105–3111.
3. Lemire, G., Ito, Y. A., Marshall, A. E., Chrestian, N., Stanley, V., Brady, L., Tarnopolsky, M., Curry, C. J., Hartley, T., Mears, W., et al. (2021). ABHD16A deficiency causes a complicated form of hereditary spastic paraplegia associated with intellectual disability and cerebral anomalies. Am. J. Hum. Genet. *108*, 2017–2023.
4. Boycott, K.M., Dyment, D.A., and Innes, A.M. (2018). Unsolved recognizable patterns of human malformation: Challenges and opportunities. Am. J. Med. Genet. Part C Semin. Med. Genet. *178*, 382-386.