Resources, challenges and way forward in rare mitochondrial diseases research [version 1; referees: 2 approved]

Neeraj Kumar Rajput¹, Vipin Singh², Anshu Bhardwaj¹

¹Open Source Drug Discovery (OSDD) Unit, Council of Scientific and Industrial Research, New Delhi, 110001, India
²Amity Institute of Biotechnology, Amity University, Noida, Uttar Pradesh, 201301, India

Abstract
Over 300 million people are affected by about 7000 rare diseases globally. There are tremendous resource limitations and challenges in driving research and drug development for rare diseases. Hence, innovative approaches are needed to identify potential solutions. This review focuses on the resources developed over the past years for analysis of genome data towards understanding disease biology especially in the context of mitochondrial diseases, given that mitochondria are central to major cellular pathways and their dysfunction leads to a broad spectrum of diseases. Platforms for collaboration of research groups, clinicians and patients and the advantages of community collaborative efforts in addressing rare diseases are also discussed. The review also describes crowdsourcing and crowdfunding efforts in rare diseases research and how the upcoming initiatives for understanding disease biology including analyses of large number of genomes are also applicable to rare diseases.

Keywords
rare disease, mitochondria, mitochondrial DNA, genome variation, crowdsourcing, crowdfunding, semantic-web, next generation sequencing

This article is included in the Rare diseases collection.
Introduction

Mitochondria are organelles present in every cell of the body (except red blood cells) and generate almost all of the energy needed by the cells to grow and sustain life. In addition to adenosine triphosphate (ATP) generation, mitochondria are involved in a large number of specialized functions in major cellular pathways including apoptosis, urea cycle, pyrimidine biosynthesis, heme synthesis, etc. The proteins that take part in these pathways are encoded both by mitochondrial DNA (mtDNA) and nuclear DNA (nDNA). mtDNA encodes only a limited number of genes (37) which code for 13 proteins, two rRNAs and 22 tRNAs. These proteins code for four respiratory complexes of the oxidative phosphorylation (OXPHOS) system. The only non-coding segment of mtDNA is the displacement loop (D-loop, 1121 bp) that contains the origin of replication of the H-strand (OH) and the promoters for L and H-strand transcription. The diseases related to mitochondrial dysfunction are due to mutations in both the mtDNA and nDNA encoded components. Genetically, mitochondrial diseases are characterized as (i) those with sporadic or maternally inherited mtDNA mutations, (ii) those with abnormalities with Mendelian transmission of the trait, i.e., disorders believed to be due to mutations in nuclear genes that control mitochondrial biogenesis, and (iii) those that are caused by nuclear genes but are misinterpreted as mitochondrial based on the biochemical findings. mtDNA point mutations which can either be maternally inherited or generated somatically have been associated with many diseases like A3243G for MELAS, A8344G for MERRF, T8993G for NARP, etc. Also, there appears to be a class of slightly deleterious mutations that modify the risks of developing certain complex diseases or trait.

Human mtDNA and nDNA mutations causing mitochondrial dysfunction are implicated in a broad spectrum of diseases affecting various tissues like brain, heart, liver, skeletal muscles, etc. The clinical symptoms of the disease depend on the cell type affected and range from loss of motor control, muscle weakness, cardiac disease to visual or hearing loss, etc. Given that mitochondria are involved in a large number of cellular pathways, it is always challenging to correlate the exact causative role of genome variation with the observed phenotype. Additionally, the varying spectrum of disease symptoms is a major deterrent in early disease diagnosis. This is clear from the fact that with over 5000 mtDNA variation reported across databases, pathogenicity assignments for most of the variation is only limited to the association with the phenotype without any conclusive evidence on its causative role. While there are a large number of resources available on various aspects of human mtDNA, a major bottleneck is the lack of documentation of genomic variation data across populations with clinical details to evaluate these variations for disease association. The need of the hour is to curate these resources using standards for data exchange over the web as also using standard ontologies for data analysis across platforms.

Most of the mitochondrial diseases have rare occurrence in the population and hence are termed as ‘Rare diseases’. As of now there are nearly 7000 rare diseases reported worldwide arising from mutations in either nDNA or mtDNA. It is estimated that nearly 300 million people in the world are affected by rare diseases. In India alone there are nearly 70 million people diagnosed with rare diseases (http://www.rarediseasesindia.org/), despite the fact that there are no standard diagnostic tests available for most rare diseases. In addition, the definition of rare diseases vary, for e.g., the European Union considers diseases to be rare when they affect no more than 5 in 10,000 people, while the Unites States of America (USA) consider a disease to be rare when affecting fewer than 200,000 people. In Asia, the threshold is 1 in 10,000. However, the prevalence of most mitochondrial diseases is not known.

This review elaborates on the available resources, the bottlenecks in rare disease research and proposes innovative ways to address these challenges. There are many reviews discussing the challenges involved in establishing genotype-phenotype correlations with mitochondrial dysfunction given its involvement in multiple pathways in a spatio-temporal manner. Here we review our current understanding of the rare disease initiatives and efforts that are ongoing globally along with specific focus on mitochondrial disease resources.

Resources and challenges: ongoing efforts & translational pipelines

The first human DNA to be completely sequenced was the human mitochondrial DNA (16569 base pairs) in 1981. Given the relatively small size and absence of repeats, sequencing and assembling the mitochondrial human genome was not as challenging and difficult as sequencing the human nuclear genome using the Sanger Sequencing technology. However, mitochondrial genome sequencing has its own unique problems given the high mutation rate and high levels of heteroplasmy. Precise determination of the levels of heteroplasmy is crucial since the level of heteroplasmy determines both the penetrance and severity of expression of some mitochondrial diseases. The next generation sequencing (NGS) technologies like Virtual terminator sequencing (Illumina), Pyrosequencing (Roche) and SOLID have allowed to overcome these limitations by providing massive parallelization, high coverage, high accuracy as compared to Sanger sequencing. Specific protocols, including long range PCR with mitochondria specific primers, and algorithms for reference based and de novo assembly have been developed to sequence mitochondrial DNA using NGS technologies. NGS-based clinical targeted gene assay for the mitochondrial genome and 108 selected nuclear genes associated with mitochondrial disorders have also been designed to facilitate the analysis and understanding of nuclear and mitochondrial variations in mitochondrial diseases. These emerging technologies offer an excellent opportunity to further dissect the molecular basis of disease manifestation. With an increasing number of individuals that may be genetically screened across different populations, excellent datasets may be available to explore the genetic basis of disease. The genomics data along with clinical and biochemical profiles may also be used to identify disease biomarkers with high sensitivity and specificity, which is a major challenge in diagnosing mitochondrial dysfunction.

Over the years, a large number of web-based resources have been developed on various aspects of mitochondrial diseases, most of them focusing on the data from mtDNA. Some of these include, MitoMap, a database on human mitochondrial variation, MitoLSDB, the largest curated data on mtDNA variation with phenotype using LOVD, MitoCarta, a resource on mitochondrial
proteins based on localization\textsuperscript{17}, MitoMiner, a mitochondrial protein identification system based on multiple evidences\textsuperscript{18,19}. MitoBreak, a curated dataset on mtDNA rearrangements\textsuperscript{20}, HmtdB, an online resource for data on mitochondrial genome sequences annotated with population and variation data\textsuperscript{21}, Mitochondrial Database (mitoDB), the mitochondrial database on clinical features seen in mitochondrial diseases\textsuperscript{22}, to name a few. Analysis pipelines and platforms have also been developed, including the MiSNPscore which assesses the role of variation in context of disease association using a combined evidence approach\textsuperscript{23}, Mit-o-matic, an analysis pipeline for clinical evaluation of mitochondrial variations from the NCATS datasets\textsuperscript{24}. More recently, the United Mitochondria Disease Foundation (UMDF) (http://www.umdf.org/site/c.8qK0J0MvF7LG/ b.7929671/k.BDF0/Home.htm) along with the National Institute of Child Health and Human Development (NICHD) (http://www.nichd.nih.gov/Pages/index.aspx) launched the Mitochondrial Disease Sequence Data Resource (MSeqDR) Consortium. The goals of this consortium is to facilitate deposition, curation, annotation and integrated analysis of genomic data for mitochondrial diseases for clinical and research communities\textsuperscript{25}. The list of various mitochondrial resources may be seen in Table 1.

Globally, attempts have also been made to systematically address the problem of rare diseases by establishing focused programs and consortia-based approaches. Therapeutics of Rare and Neglected Diseases (TRND) (http://www.ncats.nih.gov/research/rare-diseases/trnd/trnd.html), a program led by the National Center for Advancing Translational Sciences (NCATS) (http://www(ncats.nih.gov/), supports the development of potential treatments for rare and neglected diseases to first-in-human trials. This approach provides a de-risking strategy making the downstream development efforts commercially viable. TRND also supports the pre-clinical studies including medicinal chemistry optimization, drug metabolism and pharmacokinetics, toxicity formulation, and others studies required to file Investigational New Drug (IND) application for regulatory approvals. The other initiatives by NCATS for rare diseases include Office of Rare Disease Research (ORDR) (http://www.orphadata.org/cgi-bin/inc/ordo_orphanet.inc.php) which coordinates a large number of collaborative research efforts towards rare diseases including support to institutes and centers, managing patient registry, human bio-specimen repository, to name a few major activities. Rare Diseases Clinical Research Network (RDCRN) (http://rarediseases.info.nih.gov/research/pages/41/rare-diseases-clinical-research-network), from ORDR, focuses on advancing medical enrollment on rare diseases by facilitating collaboration, study enrollment and data sharing. It also connects scientists from multiple disciplines across various clinical sites globally to work with patient advocacy groups. The North American Mitochondrial Disease Consortium (NAMDC) (http://www.rarediseasesnetwork.org/namdc/), a part of RDCRN, specially works towards collecting information from mitochondrial disease patients in a clinical patient registry. In addition to periodically updating the patients on mitochondrial diseases, NAMDC also helps researchers to identify and recruit patients for future studies. Data generated as part of the various initiatives involving patient information are managed by the NCATS Global Research Patient Registry Data Repository (GRDR- http://www.ncats.nih.gov/research/rare-diseases/grdr/grdr. html). This is a web-based resource that aggregates de-identified patient information across many registries and provides a Glimpse Unique Identifier (GUID) to each patient data. GUID allows for patient follow-up across different registries, diseases, studies and countries and also ensures that clinical information is also mapped to bio-specimen datasets. Bridging interventional Development Gaps (BrIDs) (http://www.ncats.nih.gov/research/rare-diseases/bridges/bridges.html) is a division of NCATS for pre-clinical innovation towards the development of new therapeutic agents both for common and rare diseases. A recent perspective on the NCATS TRND and BrIDGs programs highlights the role of team effort where academia, biotech and pharma industries, patient communities, advocacy groups, regulators, and government support, all are needed to navigate through the translational Valley of Death\textsuperscript{26}. National Organization for Rare Disorders (NORD) (http://www.rarediseases.org/) is a non-profit organization with the aim to improve the lives of all people affected by rare diseases. The services offered by NORD include identification, treatment and cure of rare disorders through advocacy, research and programs of education. Similarly, Global genes is a non-profit organization for patient advocacy and work towards building awareness and providing connections and resources for rare disease patients (http://globalgenes.org/). Similar efforts in the European subcontinent have led to the establishment of Orphanet, a consortium of 40 countries coordinated by the French INSERM team and which hosts a reference portal for rare diseases and orphan drugs. Orphanet hosts a directory of information of expert clinics, medical labs, clinical trials, patient organizations, etc. (http://www.orpha.net/consor/cgi-bin/index.php). The joint effort of the European Commission and the NIH established IRDiRC, the International Rare Diseases Research Consortium, in 2011. This international consortium of researchers and organizations aims to deliver 200 therapies and means to diagnose most rare diseases by 2020, and as per their reports, the targets are being delivered (http://www.irdr.org/). The list of various rare diseases resources may be seen in Table 1.

As mentioned earlier, there are no clear estimates of the number of mitochondrial rare diseases. In order to get an approximation, the list of 6537 rare diseases was taken from Global genes website. To find out which rare disease is also a mitochondrial disease, a list of mitochondrial diseases was taken from the Mitochondrial Database (mitoDB)\textsuperscript{(51)}\textsuperscript{2}. In addition, information on mitochondrial diseases is also referred from UMDF (41) and MITOMAP (51). On comparing rare disease list with the mitochondrial diseases lists, 18 rare mitochondrial diseases were identified. It is important to mention here that since an exact match was performed we could have missed diseases for which abbreviation or synonyms are used. This also highlights that standard ontologies and terms are not systematically followed making data intractable for automated data analysis. We further checked the prevalence of these 18 rare diseases from the Rare Diseases India (http://www.rarediseasesindia.org/) and Orphanet portals. As may be seen on Table 2, data are available for few diseases only.

**Table 1**

| Unique Identifier (GUID) to each patient data. GUID allows for patient follow-up across different registries, diseases, studies and countries and also ensures that clinical information is also mapped to bio-specimen datasets. Bridging interventional Development Gaps (BrIDs) (http://www.ncats.nih.gov/research/rare-diseases/bridges/bridges.html) is a division of NCATS for pre-clinical innovation towards the development of new therapeutic agents both for common and rare diseases. A recent perspective on the NCATS TRND and BrIDGs programs highlights the role of team effort where academia, biotech and pharma industries, patient communities, advocacy groups, regulators, and government support, all are needed to navigate through the translational Valley of Death\textsuperscript{26}. National Organization for Rare Disorders (NORD) (http://www.rarediseases.org/) is a non-profit organization with the aim to improve the lives of all people affected by rare diseases. The services offered by NORD include identification, treatment and cure of rare disorders through advocacy, research and programs of education. Similarly, Global genes is a non-profit organization for patient advocacy and work towards building awareness and providing connections and resources for rare disease patients (http://globalgenes.org/). Similar efforts in the European subcontinent have led to the establishment of Orphanet, a consortium of 40 countries coordinated by the French INSERM team and which hosts a reference portal for rare diseases and orphan drugs. Orphanet hosts a directory of information of expert clinics, medical labs, clinical trials, patient organizations, etc. (http://www.orpha.net/consor/cgi-bin/index.php). The joint effort of the European Commission and the NIH established IRDiRC, the International Rare Diseases Research Consortium, in 2011. This international consortium of researchers and organizations aims to deliver 200 therapies and means to diagnose most rare diseases by 2020, and as per their reports, the targets are being delivered (http://www.irdr.org/). The list of various rare diseases resources may be seen in Table 1.

As mentioned earlier, there are no clear estimates of the number of mitochondrial rare diseases. In order to get an approximation, the list of 6537 rare diseases was taken from Global genes website. To find out which rare disease is also a mitochondrial disease, a list of mitochondrial diseases was taken from the Mitochondrial Database (mitoDB)\textsuperscript{(51)}\textsuperscript{2}. In addition, information on mitochondrial diseases is also referred from UMDF (41) and MITOMAP (51). On comparing rare disease list with the mitochondrial diseases lists, 18 rare mitochondrial diseases were identified. It is important to mention here that since an exact match was performed we could have missed diseases for which abbreviation or synonyms are used. This also highlights that standard ontologies and terms are not systematically followed making data intractable for automated data analysis. We further checked the prevalence of these 18 rare diseases from the Rare Diseases India (http://www.rarediseasesindia.org/) and Orphanet portals. As may be seen on Table 2, data are available for few diseases only.

**Crowdsourcing and crowdfunding rare diseases research and development**

It is evident from a simple search in the clinical trials registry (https://clinicaltrials.gov/) that only 0.2% of clinical trials ever done, ongoing, terminated or planned are for rare diseases. This clearly indicates that the major challenge for rare diseases research is the cost involved in research and development given the poor...
**Table 1. Lists of mitochondrial and rare diseases resources.** There are a large number of resources developed both for mitochondrial community and rare diseases community. The first section lists the resources available on mitochondrial resources including Support and Advocacy groups, databases and analysis pipelines, research and patient networks. The second section lists resources on rare diseases.

<table>
<thead>
<tr>
<th>S. No</th>
<th>Name</th>
<th>Description/URL</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mitochondria Diseases</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>a. Support and Advocacy</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.</td>
<td>International mito-patients (IMP)</td>
<td>IMP is a network of patient organization working on mitochondrial disease.</td>
</tr>
<tr>
<td></td>
<td></td>
<td><a href="http://www.mitopatients.org/index.html">http://www.mitopatients.org/index.html</a></td>
</tr>
<tr>
<td>2.</td>
<td>MitoAction</td>
<td>MitoAction is a support group which provides education and advocacy for mitochondrial diseases patients.</td>
</tr>
<tr>
<td>3.</td>
<td>United Mitochondrial Disease Foundation (UMDF)</td>
<td>Promotes research and education for the diagnosis, treatment, and cure of mitochondrial disorders and to provide support to affected individuals and families.</td>
</tr>
<tr>
<td></td>
<td></td>
<td><a href="http://www.umdf.org">http://www.umdf.org</a></td>
</tr>
<tr>
<td>4.</td>
<td>The Children's Mitochondrial Disease Network</td>
<td>The Children's Mitochondrial Disease Network is an information and support group for patients with mitochondrial disorders.</td>
</tr>
<tr>
<td></td>
<td></td>
<td><a href="http://www.emdn-mitonet.co.uk/index.htm">http://www.emdn-mitonet.co.uk/index.htm</a></td>
</tr>
<tr>
<td><strong>b. Research</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.</td>
<td>Mitochondrial Research Guild (MRG)</td>
<td>MRG is an organization of Seattle Children's Hospital that works to raise awareness, promoting research and improving quality of medical care.</td>
</tr>
<tr>
<td>7.</td>
<td>The Mitochondrial Medicine Society (Mitosoc)</td>
<td>Mitosoc is a society for physicians, clinicians and researchers working for the better diagnosis, management and treatment of mitochondrial diseases.</td>
</tr>
<tr>
<td></td>
<td></td>
<td><a href="http://mitosoc.org">http://mitosoc.org</a></td>
</tr>
<tr>
<td>8.</td>
<td>The Mitochondrial Research Society (MRS)</td>
<td>Organization for scientist and researchers working for the cure of mitochondrial diseases.</td>
</tr>
<tr>
<td></td>
<td></td>
<td><a href="http://www.mitoresearch.org/index.html">http://www.mitoresearch.org/index.html</a></td>
</tr>
<tr>
<td>9.</td>
<td>The Rare Mitochondrial Disease Service for Adults and Children (RMDSAC)</td>
<td>RMDSAC is an information web resource for all aspects of mitochondrial disease.</td>
</tr>
<tr>
<td></td>
<td></td>
<td><a href="http://www.mitochondrialncg.nhs.uk">http://www.mitochondrialncg.nhs.uk</a></td>
</tr>
<tr>
<td><strong>c. Database and Analysis Pipelines</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10.</td>
<td>Human Mitochondrial Database (HmtDB)</td>
<td>An online web resource for population genetics and mitochondrial disease.</td>
</tr>
<tr>
<td></td>
<td></td>
<td><a href="http://www.hmtdb.uniba.it/hmdb/">http://www.hmtdb.uniba.it/hmdb/</a></td>
</tr>
<tr>
<td>11.</td>
<td>MitoBreak</td>
<td>MitoBreak is an online resource of curated datasets of mtDNA breakpoints.</td>
</tr>
<tr>
<td></td>
<td></td>
<td><a href="http://mitobreak.portugene.com/cgi-bin/Mitobreak_home.cgi">http://mitobreak.portugene.com/cgi-bin/Mitobreak_home.cgi</a></td>
</tr>
<tr>
<td>12.</td>
<td>MitoCarta</td>
<td>MitoCarta is Data inventory for genes for Human and Mouse which encodes proteins with strong support for mitochondrial localization.</td>
</tr>
<tr>
<td></td>
<td></td>
<td><a href="http://www.broadinstitute.org/pubs/MitoCarta/">http://www.broadinstitute.org/pubs/MitoCarta/</a></td>
</tr>
<tr>
<td>13.</td>
<td>Mitochondrial Disease Database (MitoDB)</td>
<td>MitoDB is a database and analysis resource for mitochondrial disease.</td>
</tr>
<tr>
<td>14.</td>
<td>Mitochondrial Disease Sequence Data Resource (MSeqDR)</td>
<td>MseqDR is a consortium for data capturing, integration, visualization and analysis of genomic data for mitochondrial diseases.</td>
</tr>
<tr>
<td></td>
<td></td>
<td><a href="https://mseqdr.org/index.php">https://mseqdr.org/index.php</a></td>
</tr>
<tr>
<td>15.</td>
<td>MitoLSDB</td>
<td>MitoLSDB is the most comprehensive web resource for mtDNA variations on LOVD platform.</td>
</tr>
<tr>
<td></td>
<td></td>
<td><a href="https://ab-openlab.csir.res.in/mitolsdb/home.php">https://ab-openlab.csir.res.in/mitolsdb/home.php</a></td>
</tr>
<tr>
<td>16.</td>
<td>MITOMAP</td>
<td>Mitomap is a database of human mitochondrial genome.</td>
</tr>
<tr>
<td></td>
<td></td>
<td><a href="http://mitomap.org/MITOMAP">http://mitomap.org/MITOMAP</a></td>
</tr>
<tr>
<td>17.</td>
<td>Mit-o-matic</td>
<td>It is an analysis tool for finding clinical correlation of human MtDNA variations.</td>
</tr>
<tr>
<td></td>
<td></td>
<td><a href="http://genome.igib.res.in/mitomatic/index.html">http://genome.igib.res.in/mitomatic/index.html</a></td>
</tr>
<tr>
<td></td>
<td></td>
<td><a href="http://mitominer.mrc-mbu.cam.ac.uk/release-3.1/begin.do">http://mitominer.mrc-mbu.cam.ac.uk/release-3.1/begin.do</a></td>
</tr>
<tr>
<td>19.</td>
<td>MitoVariome</td>
<td>MitoVariome is a database of human mitochondrial DNA variation.</td>
</tr>
<tr>
<td></td>
<td></td>
<td><a href="http://variome.kobic.re.kr/MitoVariome/index.jsp">http://variome.kobic.re.kr/MitoVariome/index.jsp</a></td>
</tr>
<tr>
<td>20.</td>
<td>MtSNPscore</td>
<td>MtSNPscore is combined evidence based approach for prioritizing disease-associated variants.</td>
</tr>
<tr>
<td></td>
<td></td>
<td><a href="http://ab-openlab.csir.res.in/snpcore/">http://ab-openlab.csir.res.in/snpcore/</a></td>
</tr>
</tbody>
</table>
return on investment. It is also well known that clinical trials are prohibitively expensive even for common diseases where finding expert clinicians and acceptable number of patients is not as challenging as for rare diseases. These issues are addressed by de-risking research, developing platforms for sharing patient data, generating centralized patient registries and offering various incentives for making the development commercially viable. All these aspects have been discussed earlier. Here we discuss the various models that have been attempted to bridge the last mile of developing novel therapeutics.

One of the major challenges in the treatment of rare diseases is clinical trials patient recruitment. To overcome this challenge, patient advocacy groups and websites like PatientsLikeMe are turning out to be a major game changer. PatientsLikeMe is a patient powered research network that allows people to connect and share their experience with other people having the same disease or condition. PatientsLikeMe started its first online community for Amyotrophic Lateral Sclerosis (ALS) in 2006 where participants could ask specific questions about the treatment options and what to expect from fellow users (http://www.patientslikeme.com/). In addition, the patients also got involved in experimenting with drugs that have not received regulatory approval. Thus data generated by these self-reporting platforms can be used to establish the efficacy and safety of a compound for rare diseases. The self-reported data help provide evidence to support or refute treatment outcomes. Despite several limitations of self-reported data, like unmeasured covariates, data reliability and controlled settings, these approaches are promising as has been shown with the quality of data gathered with smartphone games which were comparable to data obtained from controlled laboratory environments. In 2011, the company expanded its scope and allowed any patient with any condition to join the community. Currently there are more than 300,000 members registered on the PatientLikeMe community with more than 2000

<table>
<thead>
<tr>
<th>S. No</th>
<th>Name</th>
<th>Description/URL</th>
</tr>
</thead>
<tbody>
<tr>
<td>21.</td>
<td>Matchmaker Exchange</td>
<td>A federated platform to facilitate matching similar phenotypic and genotypic profiles (matchmaking) through standardized application programming interfaces (APIs) <a href="http://matchmakerexchange.org/">http://matchmakerexchange.org/</a></td>
</tr>
<tr>
<td>22.</td>
<td>Beacon Project</td>
<td>A web service for comparing genomic data maintaining data privacy <a href="http://ga4gh.org/#/beacon">http://ga4gh.org/#/beacon</a></td>
</tr>
<tr>
<td>23.</td>
<td>Global Alliance for Genomics and Health (GA4GH)</td>
<td>GA4GH is a coalition for effective and responsible data sharing for advancement in human health through genomics medicine. <a href="http://genomicsandhealth.org/">http://genomicsandhealth.org/</a></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Rare Diseases</th>
<th>Name</th>
<th>Description/URL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Genetic and Rare Disorder Information Center (GARD)</td>
<td>GARD is an information center for genetic and rare diseases. <a href="http://rarediseases.info.nih.gov/gard">http://rarediseases.info.nih.gov/gard</a></td>
</tr>
<tr>
<td>3.</td>
<td>Global Rare Disease Patient Registry Data Repository (GRDR)</td>
<td>GRDR is a global data repository for patient information, clinical data and clinical trials data for researchers. <a href="https://grdr.ncats.nih.gov/">https://grdr.ncats.nih.gov/</a></td>
</tr>
<tr>
<td>4.</td>
<td>International Rare Disease Research Consortium (IRDiRC)</td>
<td>IRDiRC is a consortium of researchers and organization working on diagnosis and therapies for rare diseases. <a href="http://www.irdirc.org/">http://www.irdirc.org/</a></td>
</tr>
<tr>
<td>5.</td>
<td>National Organization for Rare Disorders (NORD)</td>
<td>A non-profit organization offering services for identification, treatment and cure of rare disorders through advocacy, research and programs of education <a href="http://www.rarediseases.org/">http://www.rarediseases.org/</a></td>
</tr>
<tr>
<td>7.</td>
<td>Patients Like Me</td>
<td>A for-profit company providing platform for patients to share their condition and experience for the benefit of the larger patient population. <a href="http://www.patientslikeme.com/">http://www.patientslikeme.com/</a></td>
</tr>
<tr>
<td>8.</td>
<td>Rare Diseases Clinical Research Network (RDCRN)</td>
<td>Focuses on advancing medical research on rare diseases by facilitating collaboration, study enrollment and data sharing <a href="https://www.rarediseasesnetwork.org/">https://www.rarediseasesnetwork.org/</a></td>
</tr>
<tr>
<td>9.</td>
<td>Rare Diseases India</td>
<td>Rare Diseases India is knowledge portal driven by volunteers, students, patients, health care professionals and experts. <a href="http://www.rarediseasesindia.org/about">http://www.rarediseasesindia.org/about</a></td>
</tr>
<tr>
<td>10.</td>
<td>The Office of Rare Diseases Research (ORDR)</td>
<td>ORDR is part of NCATS that supports research and provides research on rare disease. <a href="http://rarediseases.info.nih.gov/">http://rarediseases.info.nih.gov/</a></td>
</tr>
</tbody>
</table>
health conditions ([http://en.wikipedia.org/wiki/PatientsLikeMe](http://en.wikipedia.org/wiki/PatientsLikeMe)). Besides, interactions of the patient advocates from the rare diseases group with major regulatory bodies has led to expedite review and approval process for rare disease treatments ([http://www.forbes.com/sites/medidata/2014/09/25/rare-disease-patient-voices-bring-change-to-the-clinical-trials-process](http://www.forbes.com/sites/medidata/2014/09/25/rare-disease-patient-voices-bring-change-to-the-clinical-trials-process)). Using another interesting strategy, Pfizer Inc. used a web-based interactive platform to evaluate the efficacy of a drug (tolterodine tartrate extended release capsules) to treat overacting bladder. This project called REMOTE is a phase IV trial under an Investigational New Drug (IND) application. The participants for the trial were recruited through an interactive web-based platform from one clinical site overseen by physician. Despite poor patient participation, the study reports that the trial outcomes are consistent with the results from the conventional trial. The observations from REMOTE is a learning experience for the trial community on the challenges involved at various levels from patient understanding to technical issues. These learnings are used to conduct another trial in Europe, REMOTE2.0, to overcome the bottlenecks faced with the first attempt. These initiatives are of immense significance in establishing the robustness of trial strategies where patient recruitment is a challenge. These mobile technologies also allow patients with complex disabilities to participate in trials. The mobile technologies in form of wearable sensors also make the trial monitoring and patient diagnosis at different time points more affordable. Subsequent efforts to share patient data also have a positive impact on patient recruitment.

It is also proposed that the N-of-1 trial method may be used to evaluate new treatments. Under this model a single patient is the entire trial and the treatment outcome is measured at different time points for pre-treatment, treatment duration and post-treatment. If the treatment outcome is measurable and the patient is cured, the treatment may be considered effective. Thus, the crowdsourcing strategy is an attractive tool for patient recruitment as it is affordable and time effective. Likewise, crowdfunding is also proposed as a viable option for financing various aspects of rare diseases research. Crowdfunding initiatives not only generate awareness about rare diseases but also lead to funding for supporting research activities. One such example is the ALS Ice Bucket Challenge leading to 1.2 million videos on Facebook and 2.2 million tweets in a matter of 3 months, which raised over $100 million of donations ([http://en.wikipedia.org/wiki/Ice_Bucket_Challenge](http://en.wikipedia.org/wiki/Ice_Bucket_Challenge)). Through innovative ideas like ALS Ice Bucket Challenge it is possible to raise funds and also increase public awareness about the rare diseases. The Rare Genomics Institute (RGI) ([http://raregenomics.org](http://raregenomics.org)) is an international non-profit providing expert network and an online crowdfunding mechanism to assist families pursue personalized research projects for diseases which would not be studied otherwise. RGI offers sequencing facilities for diagnosis, expert guidance on sequencing and systematic interpretation of the data generated. A recently published success story is that of a 3-year old girl showing symptoms of involuntary eye movement, small-sized head, involuntary muscle contraction, development delay and progressive decline. Whole exome sequencing of the family

### Table 2. List of rare mitochondrial diseases

The table provides a list of diseases that cause mitochondrial dysfunction and are also reported rare diseases.

<table>
<thead>
<tr>
<th>Rare Diseases</th>
<th>India Per 100,000</th>
<th>Number of reported cases</th>
<th>Europe Per 100,000</th>
<th>Number of reported cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acute megakaryoblastic leukemia</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Barth syndrome</td>
<td>NA</td>
<td>0.22</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Chronic progressive external ophthalmoplegia</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Coenzyme Q10 deficiency</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Familial bilateral striatal necrosis</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Friedreich ataxia</td>
<td>NA</td>
<td>2</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Infantile onset spinocerebellar ataxia</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>24</td>
</tr>
<tr>
<td>Kearns Sayre syndrome</td>
<td>NA</td>
<td>2</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Leber hereditary optic neuropathy</td>
<td>7</td>
<td>84000</td>
<td>1.5</td>
<td>NA</td>
</tr>
<tr>
<td>Leig syndrome</td>
<td>3</td>
<td>36000</td>
<td>2.75**</td>
<td>NA</td>
</tr>
<tr>
<td>Maternally inherited Leigh syndrome</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Pyruvate carboxylase deficiency</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Pyruvate dehydrogenase deficiency</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Pyruvate dehydrogenase phosphatase deficiency</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Rett syndrome</td>
<td>4</td>
<td>48000</td>
<td>4</td>
<td>NA</td>
</tr>
<tr>
<td>Spastic paraplegia 7</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Spinocerebellar ataxia 28</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Sudden infant death syndrome</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

** indicates birth prevalence
revealed a novel mutation that causes mental retardation and severe
developmental delays. Through the RGI platform, $5000 was raised
in 50 days to carry out the genome sequencing in the girl and her
family, which led to identification of the novel mutation. These
examples underline the potential of crowdfunding in addressing
scientifically challenging and socially important problems.

With the increasing disease burden, the need of the hour is to estab-
lish new models for ensuring that clinical trials are sustainable and
transparent. Clinical trials are generally demanding on both human
and other resources. It is not possible to apply the conventional
methods to rare disease trials given the sheer costs involved. To
overcome these bottlenecks, efforts are being made lately to reverse
the approach of rare disease patient recruitments for clinical trials
by taking the trials to the patients and not vice versa. This approach
demands establishing patient-centric sites which is again challeng-
ing given their extremely sparse distribution. Telemetry innovations
may offer solutions to this problem, which allows health monitor-
ing of the patients remotely. This model is not just suitable for rare
diseases but is also applicable to common diseases trials given that
many clinical trial sites never recruit patients (http://www.clini-
calleader.com/doc/how-rare-disease-know-how-can-shape-big-
pharma-clinical-trials-0001).

Open data sharing and analysis platforms as drivers for innovation

Open access platforms which allow for data integration and
exchange of ideas to facilitate the process of bringing new therapeu-
tic interventions and diagnostic methods to the market are needed to
drive sustained research and development for rare diseases. It is also
imperative to synchronize the efforts at all levels of research, diag-
nosis and regulatory guidelines for evidence-based clinical deci-
sion-making. As part of the platform, it is also important to involve
regulatory agencies from very early on in the discovery pipeline so
that the progression through development and approvals is quicker
and more affordable, which is the bottleneck with rare diseases.
One strategy which seems to work well is to apply approaches of
drug repurposing where it is crucial to understand the mechanism
of disease manifestation and subsequently applying various mod-
eling approaches to find new indications of existing drugs in the
market (http://www.fda.gov/ForIndustry/DevelopingProductsfor-
RareDiseasesConditions/HowtoapplyforOrphanProductDesigna-
tion/ucm216147.htm). This strategy will be even more successful
when one can systematically list the phenotypic parameters pro-
vided by patient groups in close collaboration with clinicians.

There is a need for innovation in technology that allows for interac-
ting the different stakeholders, namely, researchers, clinicians and
patient groups. Such resources will go a long way in future to allow
patients to perform diagnostic procedures with greater accuracy
and help clinicians identify the best possible route for therapeu-
tic interventions. This is only possible if the participating members
use standard terms to share the data. Various ontologies have been
developed over years for capturing function of genes like the Gene
Ontology, PAGE-OM and VariO for capturing features of vari-
ation, UMLS which includes MeSH, RxNORM and SNOMED
CT for capturing clinical details and Human Phenotype Ontology for
phenotype data, to name a few. These ontologies are meant to

function as a common set of vocabulary that needs to be shared on
a community collaborative platform so that the semantic network
of these terms allows for deciphering and interpreting novel patterns
from these networks. Such a semantic-based platform will be ame-
nable for plugging in data and new methods with increasing under-
standing of disease biology. It will also ensure that the data generated
as part of negative results are also shared systematically with scien-
tific community. Orphanet Rare Disease Ontology (ORDO) (http://
www.orphadata.org/cgi-bin/inc/ordo_orphanet.inc.php) is an effort
to achieve these goals. ORDO offers definitions, classification of
rare diseases, gene-disease relations, epidemiological data and con-
nections with other terminologies like MeSH, UMLS (http://www.
.nlm.nih.gov/research/umls/), OMIM, UniProtKB, HGNC, ensembl,
Reactome, etc. It is imperative that existing resources utilize these ontologies for data sharing allowing for data interoper-
ability across different platforms. Figure 1 illustrates the scope
of the proposed integrative platform describing the scientific chal-
enges involved in establishing genotype-phenotype correlations
and how the existing resources and community collaborations may
be converged towards a systems level understanding of the disease
biology.

What’s new?

‘Internet of DNA’ is listed as one of the top 15 breakthrough tech-
nologies of 2015 (http://technologyreview.com/featuredstory/
535016/Internet-of-DNA/). The technology needed to harness the
power of genomics lies in comparing the genetic information from
a large number of individuals with medical records. This currently
is a huge challenge, partly because of the technical reasons of mov-
ing petabytes of data across different labs, but especially due to the
privacy issues surrounding patient information. Both these issues
should be addressed to ensure that the ever-increasing amount of
genomic and clinical data piling up in laboratories and hospitals
are utilized optimally. Upcoming initiatives like the MatchMaker
Exchange are aiming to bring the genotype and phenotype data
together on a common platform (http://matchmakerexchange.org/).
Global Alliance for Genomics and Health also known as GA4GH
is an organization which provides protocols, APIs and file formats
for effective and responsible sharing of genomic and clinical data
(http://genomicsandhealth.org/). The organization goal is to over-
come challenges likes ethics and privacy involved with sharing of
genomics data and to accelerate the potential of genomic medicine
for advancement of human health. As discussed before, the need of
the hour is to let the patients decide on who will access their data
and how these may be used. This is only possible if the information
generated using patient samples is made available in real-time. It
is also important that other components of this major collaborative
strategy are also part of a community platform that allows for gated
access to patient data with patients deciding how and with whom
their data should be shared.

In the changing paradigms of disease treatment, a very recent
approval is made by Britain on Mitochondrial donation (http://www.
theguardian.com/politics/2015/feb/24/uk-house-of-lords-approves-
conception-of-three-person-babies). In this approach in vitro
fertilization (IVF) technique is used with biological material coming
This section describes the challenges involved in dissecting the impact of genomic variation in disease association. As shown, mitochondria are involved in a large number of cellular processes including energy metabolism. Genes encoded by mitochondrial and nuclear genomes carry out these functions. The protein subunits of the complexes of the electron transport chain are encoded both by mtDNA (red) and nDNA (green). In order to understand the complex genotype-phenotype correlation it is imperative to identify the molecular interactions at systems level (protein-protein interaction network, metabolic network, signaling network, regulatory network). It is important to curate the information needed to generate these networks from literature and existing resources.

The resources currently available for mitochondrial dysfunction include databases, web servers and analysis pipelines as shown. For generating systems level models, these resources may be integrated systematically. There are various ongoing efforts involving researchers, clinicians and or patient groups. It is proposed that a community collaborative open access platform is a must to interface these communities. In order to establish such a platform that allows geographically different communities to work together, globally accepted ontologies in a language independent representation are needed.
from three parents, mother and father (contributing 98.8% genetic material) and a female donor (contributing 0.2% genetic material). This three-parent IVF approval in UK has received mixed reactions and only time will address the concern of the long-term implications of the same.

Conclusions

Rare diseases affect over 300 million people globally, however the true burden of these diseases on human health remains to be determined. Rare genetic variants are disease causing and lead to a personalized disease manifestation. Thus, it is time to review the disease definition considering both the molecular mechanisms involved and environmental factors leading to differential phenotypes. This will allow for a better understanding of both rare and common. On the other end, a paradigm shift in drug discovery and development is also needed to translate the effort in understanding disease mechanisms to identify potential therapeutic routes. Newer models and platforms that allow involvement of patient communities in research and development is also expected to offer solutions to patients suffering from rare diseases who may then benefit from appropriate treatment options. Community collaborative approaches for research and funding offer an unprecedented opportunity for making new discoveries and translating to therapeutic interventions.

Author contributions

AB conceived and designed the article outline. NKR and AB worked on the tables and figures. AB, NKR and VS wrote the manuscript. All authors have read and approved the manuscript.

Competing interests

No competing interests were disclosed.

Grant information

NKR acknowledges the CSIR-GENESIS project for providing research fellowship.

I confirm that the funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References


Frost JH, Massagyi MP: Social uses of personal health information within PatientsLikeMe, an online patient community: what can happen when patients have access to one another’s data. J Med Internet Res. 2008;10(3):e15. PubMed Abstract | Publisher Full Text | Free Full Text


Open Peer Review

Current Referee Status: ✔️ ✔️

Version 1

Referee Report 01 July 2015
https://doi.org/10.5256/f1000research.6657.r9177

✔️ Vanniarajan Ayyasamy
Department of Molecular Genetics, Aravind Medical Research Foundation, Madurai, Tamil Nadu, India

This article really serves as a good resource for researchers, clinicians and also patients.

I would suggest the authors to compare the available databases of the mitochondrial research and may provide a better summary.

Competing Interests: No competing interests were disclosed.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 07 Aug 2015

Anshu Bhardwaj, Open Source Drug Discovery Unit, Council of Scientific and Industrial Research, India

We thank the reviewer for encouraging comments. The scope of the current review was to provide an end-to-end view of the resources, challenges and way forward for addressing translational challenges in mitochondrial diseases. The comparison of databases of the mitochondrial research needs a separate review as the parameters for comparison need to be comprehensive and the information related to each resource should be discussed in details. The authors are working on the same and will share the details as soon as it is ready.

Competing Interests: No competing interests.

Referee Report 26 June 2015
https://doi.org/10.5256/f1000research.6657.r8964

✔️ Xin Qi
Center of Mitochondrial Disease, School of Medicine, Case Western Reserve University, Cleveland, OH, USA
This is a very good and informative review paper, which summarized nicely on the resources on research of mitochondrial diseases ranging from clinical setting to basic science.

**Competing Interests:** No competing interests were disclosed.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

---

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com