

REVIEW

Open Access

# Genomics of rare genetic diseases—experiences from India



The GUARDIAN Consortium, Sridhar Sivasubbu<sup>1\*</sup> and Vinod Scaria<sup>1\*</sup>

## Abstract

Home to a culturally heterogeneous population, India is also a melting pot of genetic diversity. The population architecture characterized by multiple endogamous groups with specific marriage patterns, including the widely prevalent practice of consanguinity, not only makes the Indian population distinct from rest of the world but also provides a unique advantage and niche to understand genetic diseases. Centuries of genetic isolation of population groups have amplified the founder effects, contributing to high prevalence of recessive alleles, which translates into genetic diseases, including rare genetic diseases in India.

Rare genetic diseases are becoming a public health concern in India because a large population size of close to a billion people would essentially translate to a huge disease burden for even the rarest of the rare diseases. Genomics-based approaches have been demonstrated to accelerate the diagnosis of rare genetic diseases and reduce the socio-economic burden. The Genomics for Understanding Rare Diseases: India Alliance Network (GUARDIAN) stands for providing genomic solutions for rare diseases in India. The consortium aims to establish a unique collaborative framework in health care planning, implementation, and delivery in the specific area of rare genetic diseases. It is a nation-wide collaborative research initiative catering to rare diseases across multiple cohorts, with over 240 clinician/scientist collaborators across 70 major medical/research centers. Within the GUARDIAN framework, clinicians refer rare disease patients, generate whole genome or exome datasets followed by computational analysis of the data for identifying the causal pathogenic variations. The outcomes of GUARDIAN are being translated as community services through a suitable platform providing low-cost diagnostic assays in India. In addition to GUARDIAN, several genomic investigations for diseased and healthy population are being undertaken in the country to solve the rare disease dilemma.

In summary, rare diseases contribute to a significant disease burden in India. Genomics-based solutions can enable accelerated diagnosis and management of rare diseases. We discuss how a collaborative research initiative such as GUARDIAN can provide a nation-wide framework to cater to the rare disease community of India.

**Keywords:** Rare disease, Genomics, India, Genetic diversity, Diagnostics, GUARDIAN, Zebrafish, iPSCs, Patient support

\* Correspondence: [sridhar@igib.in](mailto:sridhar@igib.in); [s.sivasubbu@igib.res.in](mailto:s.sivasubbu@igib.res.in); [vinods@igib.in](mailto:vinods@igib.in); [vinods@igib.res.in](mailto:vinods@igib.res.in)

<sup>1</sup>CSIR Institute of Genomics and Integrative Biology, Delhi 110025, India  
Full list of author information is available at the end of the article



© The Author(s). 2019 **Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated.

## Background

### Population architecture and genetic diversity in India

India is the sixth largest country in the world in terms of its geographical area and the second largest country in population density. The people of the country are diverse in terms of their social, linguistic, cultural, and racial backgrounds. Evolutionarily, the Indian subcontinent has been a corridor for different migratory waves arising from Africa, through land as well as coastline routes [1, 2]. Genetic studies have shown that there are four distinct ancestral groups in mainland India, and a separate ancestry in the Andaman and Nicobar islands [3, 4]. On the basis of ethno-racial grounds, the four major groups in India can be classified as the Caucasoids, Australoids, Mongoloids, and Negritos. The Indian population comprises of over 4000 anthropologically distinct groups speaking more than 300 languages [5], suggesting that linguistic stratification is highly tied to the geographical niches of each sub-population [6–10]. Further, the population is also sub-classified into tribes and castes based on cultural and social backgrounds [8]. These different layers of population stratification have led to the richness in diversity of India.

The genetic diversity is well reflected in the mitochondrial DNA (mtDNA), Y chromosomes, and candidate genes/markers, which have provided a fair understanding of the relatedness and divergence of specific communities or tribes of India [6, 8, 11–17]. The prevalence of consanguinity in marriages, due to cultural and social practices, in many sub-populations in India has led to the accumulation of genetic traits within communities [3, 18]. Studies have shown a high level of relatedness within subgroups suggesting accumulation of deleterious variations [19, 20]. These studies indicate that the ancestors of different subpopulations in India may have arisen from different waves of migration with relatively limited founding members, implying the source of genetic distinction, while regionally and culturally distinct groups continue to be genetically unique due to the practices of inbreeding.

A national genome-wide approach to understand the population architecture and look for markers specific to the Indian subcontinent was undertaken by the Indian Genome Variation (IGV) consortium, which used single-nucleotide polymorphisms (SNPs) to type 900 genes from over 1800 individuals across 55 endogamous populations. High heterozygosity values, varying allele frequencies, and common polymorphic haplotypes of sub-populations were shown to underline the heterogeneity within the Indian population. Additionally, unique mutations were discovered within the subcontinent, with concomitant founder effects [10, 21, 22].

The findings of the IGV consortium have led to the identification of specific markers and better understanding of genotype-phenotype correlations in Indian sub-

populations. The phenotypically distinct outcomes of sub-population specific genotypes could be shown in susceptibility or resistance towards *Plasmodium falciparum* [23–27], risk of contracting glaucoma [28], homocysteine levels [29], and risk of developing high-altitude pulmonary edema [30, 31], among other examples. Further, case-control studies in ethnically matched groups as defined by IGV consortium allowed identification of Indian-specific susceptibility markers in genes causing Parkinson's disease, Wilson disease, and albinism [32–35]. Sub-population-specific responses to various drugs have also been documented, based on differences in the allele frequencies of variants in metabolizer enzyme genes, across various ethnicities in India [36–38].

Thus, the extensive genetic heterogeneity and the endogamous cultural practices clearly suggest that there is a need to demarcate genetic affinities and distinctions among sub-populations. These findings also underscore the genetic distinction of the Indian population from the populations of other countries, warning against the imputation of genetic information from other populations. Evidently, a generalization of the population architecture can lead to erroneous interpretations in clinical settings.

### Genetic diversity of India: a driver of high-genetic disease prevalence

India, being a melting pot of genetic diversity, is also home to strict inbreeding practices and founder effects, which have resulted in the accumulation of deleterious genetic variations [39]. The reported prevalence of birth defects in India is 64.4 per 1000 live births [40]. The high genetic burden in India has been highlighted by independent studies [41–44]. The lack of a national newborn screening program until recently has led to a distending proportion of the Indian population ailing with genetic diseases [45]. Inborn errors of metabolism (IEM), which is a nation-wide issue, can be addressed on being identified at the neonatal stages [46, 47]. Hemoglobinopathies including sickle cell anemia, thalassemia, pose a significant burden in India, and are known in specific sub-populations [48, 49]. Down syndrome is another genetic disorder, which is the major cause of mental retardation, with a frequency of approximately 1 in 1000 births [50]. A database for cataloging genetic diseases, the Indian Genetic Disease Database (IGDD) has been set up, version 1.0 of which housed information on variants in 63 genes corresponding to 52 genetic diseases known in the Indian population [51]. The database is freely available and currently holds information on over 100 genetic diseases from around 3500 patients [52].

What is striking, apart from the high prevalence of monogenic diseases, is the heterogeneity in the outcome of

the same disease. The clinical heterogeneity in blood disorders in India has been attributed to subpopulation-specific variations and allele frequencies [53–57]. Similarly, the phenotypic spectrum of Spinocerebellar ataxias (SCA) and their pathogenic variants have been shown across Indian subpopulations [42]. Ethnicity-dependent mitochondrial haplotypes have also been shown to give rise to differences in penetrance in the mitochondrial disease Leber's hereditary optic neuropathy (LHON) [58]. Population-specific genetic variations and susceptibility to diseases have been shown in hereditary cardiomyopathy [59, 60] and drug/toxin metabolism [61]. The genetic heterogeneity, which was thought as an advantage, is, in fact, contributing to the high prevalence of genetic diseases in India. Several studies have also shown that the genetic variations and frequency information observed in population worldwide are not fully relevant to the Indian context [62–64]. Thus, it is important to document the true extent of genetic variation and burden of genetic diseases in Indian settings.

A number of genome-scale datasets of Indians have surfaced in recent years. These include an initiative by the IGV consortium of six laboratories affiliated to the Council of Scientific and Industrial Research (CSIR) with other key players, that typed SNPs and known markers scattered among 1000 genes [10, 21, 22, 65]. This was also followed by whole-genome sequencing of Indians from the USA [66] and from India [67, 68], in addition to several large-scale projects which sequenced healthy individuals who are descendants of Indian immigrants and from specific Indian sub-populations [69–72]. Genomes of healthy individuals from different parts of India were sequenced subsequently [73–77]. These initiatives have culminated in efforts to meta-analyze and integrate datasets, which has resulted in resources such as the South Asian Genomes and Exomes (SAGE) [76] and INdian EXome database

(INDEX-db) [78]. In addition, several disease or application specific databases developed in India provide a rich source of information about the genetic diversity and underlying genetic disease prevalence in India (Table 1).

It is to be noted that given the heterogeneity shown by IGV and other studies, the number of Indian genomes and exomes that are available till date under-represents the peninsula's diversity. This gap in the availability of baseline genetic information can hence act as a barrier in understanding the causes of diseases that are prevalent in the country and calls for a nation-wide genome project, as being undertaken in other parts of the world [82].

## Main text

### Rare diseases: a significant burden for India

Rare diseases or orphan diseases are defined as those which afflict a minimal fraction of a population. An attempt to identify the parameters that can be used to define a rare disease was made by the 'Rare Disease Terminology & Definitions Used in Outcomes Research Working Group.' The study concluded that a disease with the average global prevalence of 40–50 cases per 100,000 people can be called as a rare disease [83]. The Orphan Drug Act (ODA) of 1983 [84] under the US law, which was instrumental in gathering attention towards rare diseases [85], defined a rare disease in the USA as a disease affecting fewer than 200,000 people of the total population. The council of the European Union defined a rare disease as 5 in 10,000 [86]. The rare disease prevalence for different countries thus varies. For instance, the respective rare disease prevalence numbers are 65 in 100,000 in Brazil [87], 1 in 2500 in Japan [83], and 33.2 per 100,000 in Taiwan [88].

The pervasive endogamy and founder effects in sub-populations have led to a high prevalence of autosomal

**Table 1** Details of publicly available resources that can aid in rare genetic disease research in India

S. no.	Databases/resources	Description/URL	Reference
1	SAGE	A compendium of genetic variants integrating South Asian whole genomes and exomes <a href="http://clingen.igib.res.in/sage/">http://clingen.igib.res.in/sage/</a>	[76]
2	IGVdb	A DNA variation database of the people of India available to researchers for understanding human biology with respect to disease predisposition, adverse drug reaction, population migration, etc <a href="http://www.igvdb.res.in/index.php">http://www.igvdb.res.in/index.php</a>	[22]
3	MtBrowse	Integrative genomics browser for human mitochondrial DNA hosting genomic variation data from over 5000 individuals with 22 disease phenotypes <a href="http://ab-openlab.csir.res.in/cgi-bin/gb2/gbrowse">http://ab-openlab.csir.res.in/cgi-bin/gb2/gbrowse</a>	[79]
4	mit-o-matic	A comprehensive cloud-based tool for clinical evaluation of mitochondrial genomic variations from NGS datasets <a href="http://genome.igib.res.in/mitomatic/help.html">http://genome.igib.res.in/mitomatic/help.html</a>	[80]
5.	INDEX-db	Database of genetic variations from the Indian population <a href="http://indexdb.ncbs.res.in/">http://indexdb.ncbs.res.in/</a>	[78]
6.	TMC SNPdb	First open source SNP database from whole exome data of 62 samples derived from cancer patients from India. <a href="http://www.actrec.gov.in/pi-webpages/AmitDutt/TMCSNP/TMCSNPdp.html">http://www.actrec.gov.in/pi-webpages/AmitDutt/TMCSNP/TMCSNPdp.html</a>	[81]
7.	IGDD	Indian Genetic Disease Database <a href="http://www.igdd.iicb.res.in/">http://www.igdd.iicb.res.in/</a>	[51]

recessive rare genetic diseases in India, compared to other parts of the world. While there is no appropriate standard definition to describe a rare disease in India, Indian Council of Medical Research (ICMR) has defined a disease as rare if it affects less than 1 person in 2500 individuals [89]. The Organization for Rare Diseases India (ORDI) has suggested a threshold of 1 in 5000 for defining rare diseases in India [90]. About 5000–8000 rare diseases have been documented all over the globe accounting for up to 6–8% of the global population [86]. Approximately, 40% of the rare diseases can be attributed to genetic factors [91]. These diseases together contribute to a significant number of individuals and the disease burden in a populous country such as India.

The estimation of the prevalence of rare genetic diseases across India is limited by the lack of a centralized clinical registry of patients with rare genetic diseases. However, extrapolating the numbers in the Indian scenario, the Foundation for Research on Rare Diseases and Disorders has estimated that about 70 million people are affected by rare diseases [92]. Rare diseases that have gained attention in the country include blood disorders, lysosomal storage diseases, primary immunodeficiency diseases, mitochondrial diseases, neurodegenerative diseases, and musculoskeletal diseases, among many others [89, 93]. A compilation of estimated prevalence/incidence of well-studied rare diseases in India has been included in Table 2.

Given the estimate of approximately 70 million people living with rare diseases, most of them undiagnosed, rare disease management contributes a huge burden for a developing country like India. The accurate socio-economic burden due to rare genetic diseases in India is unknown. Incidentally, the social impacts of

hemophilia have been recorded adequately, in spite of an underestimated prevalence due to lower case reporting [94]. Other studies have shown that government interventions can reduce the out-of-pocket expenditure of patients [101, 102]. A recent study showed a yearly expenditure of transfusion-dependent thalassemics attending a tertiary care center in India, to be Rs. 41,514 to 1,51,800. This is equivalent to USD 629–2300 with an average of Rs. 74,948 (USD 1135), amounting to almost 40% of the annual income of an Indian family [103]. In recent years, several initiatives have been taken by Indian organizations, both government and non-government, to address rare diseases and the availability of orphan drugs to help ailing patients [104]. However, there are several challenges including physician training, availability of molecular diagnosis, standard treatment protocols, and availability of drugs, among others, that need to be addressed to reduce the rare disease burden in India.

#### Population scale initiatives for addressing rare diseases in India

Despite over 70 million individuals being affected by rare diseases, India has limited resources committed to treating or understanding rare diseases. In recent years, Indian Council of Medical Research (ICMR) has taken a step towards bridging the gap between patients suffering from rare genetic diseases and healthcare providers by launching The Indian Rare Disease Registry. The registry acts as a common repository for data concerning rare disease patients throughout the country [105]. Furthermore, there are examples of how various organizations, both government and non-government, have developed programs for addressing the rare disease challenge in

**Table 2** List of rare genetic diseases with estimated prevalence/ incidence in India

S. no.	Rare disease	Frequency in India	Measure of estimation	State/region	Reference	Global prevalence (Orphanet)
1	Hemophilia A	0.9 per 100,000	Prevalence	All across India	[94]	1–9/100,000 (ORPHA:98878)
2	Hemophilia B	0.1 per 100,000	Prevalence	All across India	[94]	1–9/100,000 (ORPHA:98879)
3	Sickle cell anemia	2–20%	Allele frequency	All across India	[57]	1–5/10,000 (ORPHA:232)
4	Beta thalassemia trait	3–4%	Carrier Prevalence	All across India	[95]	1–9/1,000,000 (ORPHA:848)
5	Parkinson's disease	6–53/100,000	Prevalence	All across India	[96]	Unknown (ORPHA:411602)
6	Duchenne muscular dystrophy and Spinal muscular atrophy	1 in 1400 male live births	Prevalence	Tamil Nadu, South India	[97]	NA
7	Cystic fibrosis	0.40%	Gene frequency	All across India	[98]	1–9/100,000 (ORPHA:586)
8	Epilepsy	2.5–11.9/1000	Prevalence	North, South, East India	[96]	–
9	Intellectual disability	10.5/1000	Prevalence	All across India	[99]	–
10	Skeletal dysplasia	19.6 per 10,000 newborns	Incidence	Karnataka, South India	[100]	< 1/1,000,000 (ORPHA:1858)

Note: The table provides a list of prevalent rare genetic disease studies carried out in India. While there were studies for many other diseases, they have been excluded since they do not represent the actual prevalence in the general population

India. However, most of these efforts are towards specific diseases areas or are targeted to a certain sub-population. Some of the notable initiatives that cater to heterogeneous rare disease patients are highlighted in this section.

Molecular Diagnostics, Counselling, Care and Research Centre (MDCRC) is a not-for-profit charitable organization which takes a holistic approach to manage Duchenne Muscular Dystrophy (DMD) patients, mostly catering to individuals from the southern part of India (Tamil Nadu). MDCRC undertakes genetic counseling in addition to providing screening for DMD and Spinal Muscular Atrophy (SMA). A pilot study by MDCRC estimated the prevalence of DMD to be 2.4 times higher as compared to global estimates [97]. The Uttar Pradesh state government had taken the commendable initiative in the year 2009 by providing anti-hemophilic factors (AHF) free of cost at various centers in the state [106], while the Maharashtra state government has provided clotting factor concentrates (CFC) to the poor sections and emergency cases since 2012 [107]. According to the hemophilia federation of India, 69% of the country is covered by AHF support [108]. These have been successful initiatives for public health in specific rare disease settings. Institute of Medical Genetics and Genomics at the Sri Ganga Ram Hospital, Delhi provides a battery of tests for several rare diseases [109] including blood disorders, metabolic disorders, muscular dystrophies, and Down syndrome [110], among others.

Sanofi-Genzyme's India Charitable Access Program (INCAP), Shire HGT's charitable access program in partnership with Direct Relief (a non-governmental organization), and Protalix Biotherapeutics have provided access to enzyme replacement therapy for lysosomal storage diseases in India [111]. Apart from these, there are a handful of commercial companies in India that offer genetic testing for rare genetic diseases, thus aiding the rare disease diagnosis requirements. In recent years, ORDI, a non-profit non-government organization in India, is providing a platform for individual rare diseases support groups to come together. They aim to set up patient registries and work with the government to create policies that are orphan disease centered. ORDI undertakes both Indian and global initiatives, and works together with at least 15 rare disease foundations/centers [90].

The Genomics for Understanding Rare Diseases: India Alliance Network (GUARDIAN) at CSIR-Institute of Genomics and Integrative Biology (CSIR-IGIB), Delhi is a unique research initiative in India that uses the power of genomics to solve and understand rare diseases. Details about the GUARDIAN program are elaborated in the next section. Apart from those listed above, several government research laboratories, hospitals, and not-for-

profit organizations also provide specialized tests for a specific patient group or community (see Tables 3 and 4 for more details).

#### **GUARDIAN**

Completion of the human genome project and the availability of the human genome reference sequence have opened up opportunities for a new era of genomic medicine. This has a tremendous impact on diagnosis, treatment, and preventive care related to genetic diseases [112–114]. The decade after the completion of the human genome sequence has ushered in significant technological advancements [115–117]. These technologies, popularly known as Next Generation Sequencing (NGS) technologies have enabled fast sequencing of genomes at an affordable cost [118, 119]. The improvements in technology have also contributed immensely to the development of complementary methods towards extraction of biological interactions between biomolecules including the transcriptome [120–122] and epigenome [123]. In addition, the integration of personal omics data provides opportunities to view the temporal dynamics of omics profiles in an individual [124, 125]. These advances have brought in a paradigm shift in current practices of medicine. Genome sequencing has significantly impacted the understanding of genetic variants and their association with diseases. Recently, exome and genome sequencing are increasingly being used to investigate the genetic bases of diseases including both monogenic as well as complex diseases such as cancer. One of the major applications of such genomic technologies in the clinical setting is the identification and annotation of variants associated with rare genetic diseases [126–130]. A rare disease patient usually undergoes three misdiagnoses and takes up to 7 years to reach the right diagnosis [131]. With genome sequencing technologies, it is now possible to look at either the entire genome or the protein-coding regions (exomes) that may harbor deleterious variations, in a reasonable time. Given the presence of unique variations in Indian populations, absent elsewhere in the world, genomics-based solutions are the way forward to tackle the high burden of rare diseases. Identifying the causative variant(s) in rare genetic diseases would be important not only in enabling accurate diagnosis but also in counseling and genetic screening applications.

The major challenges in realizing the full potential of genomics technologies for identifying genetic disease-causing variants in India are manifold. These include the uniqueness of the Indian genetic pool, lack of a program for identifying rare genetic diseases, and a comprehensive registry of rare genetic diseases, logistics of sample procurement and processing, common protocols for genome sequencing and computational analysis, and methodologies for validating the functionality of the

**Table 3** List of major research centers working on rare diseases in India

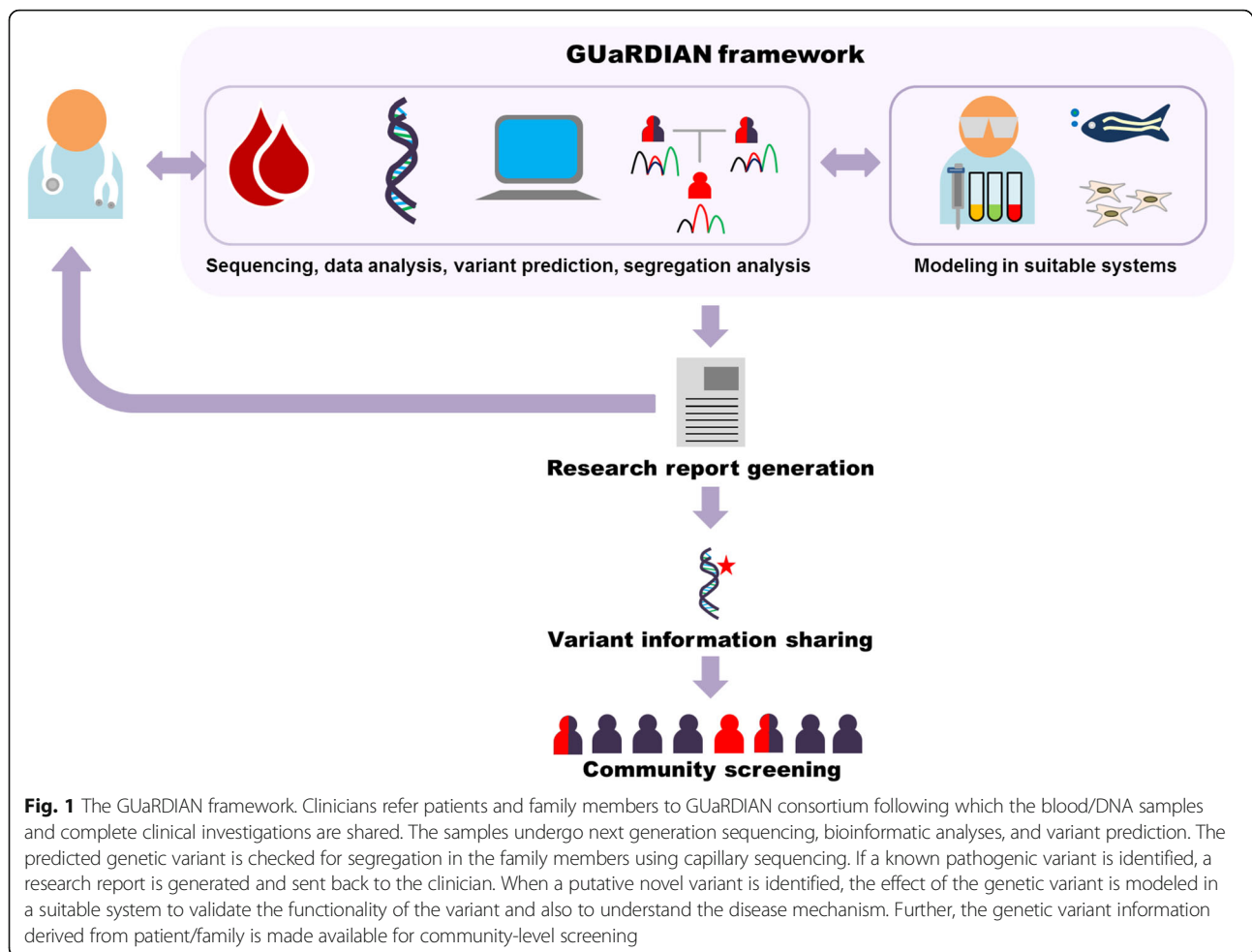
S. no.	Research centers	Major rare disease research areas
1	All India Institute of Medical Sciences (AIIMS), New Delhi	A referral hospital with multiple specialties. Major rare disease areas include skin disorders, mitochondrial disorders, neurological disorders, cardiac disorders, developmental disorders, and pediatric disorders, among others
2	Amrita Institute of Medical Sciences and Research Centre, Cochin	Lysosomal storage disorders (LSDs), Werner syndrome
3	Anthropological Survey of India (ASI), Kolkata	Hemoglobinopathies
4	CSIR Central Drug Research Institute (CDRI), Lucknow	Progressive external ophthalmoplegia
5	CSIR Centre for Cellular and Molecular Biology (CCMB), Hyderabad	Mitochondrial disorders, hemoglobinopathies, infertility
6	Centre for Human Genetics (CHG), Bengaluru	Inherited metabolic diseases (IMDs), EB, lysosomal storage disorders
7	Christian Medical College and Hospital (CMC), Vellore	A referral hospital with multiple specialties. Major rare disease areas include blood disorders among others
8	Center for Genetic Studies and Research, MMM Hospital, Chennai	Rare chromosomal diseases
9	FRIGE's Institute of Human Genetics, Ahmedabad	Hemoglobinopathies, musculopathies, neurodegenerative diseases, lysosomal storage disorders, among other genetic diseases
10	CSIR Indian Institute of Chemical Biology (IICB), Kolkata	Oculocutaneous albinism (OCA), Wilson disease (WD), autism
11	Indian Institute of Science (IISc), Bengaluru	Primary microcephaly, anencephaly, Parkinson's disease, Wilson disease, and neuromuscular disorders
12	Indira Gandhi Institute of Child Health, Bangalore	Lysosomal storage disorders, Prader-Willi syndrome, and skeletal dysplasia, among other rare diseases
13	CSIR Institute of Genomics and Integrative Biology (IGIB), New Delhi	A specialized laboratory for research in rare genetic diseases including skin disorders, ataxias, cardiac disorders, neurological disorders, primary immunodeficiency disorders, endocrinology disorders, nephrological disorders, mitochondrial disorders, Wilson disease, hemoglobinopathies, lysosomal storage disorders, and developmental disorders, among others
14	Jawaharlal Institute of Postgraduate Medical Education and Research (JIPMER), Puducherry	A referral hospital with multiple specialties. Major rare disease areas include Werner syndrome, Fanconi anemia, split hand-split feet syndrome, skin disorders, incontinentia pigmenti
15	Jawaharlal Nehru University (JNU), New Delhi	GNE myopathy
16	JK Lone Hospital, SMS Medical College, Jaipur	A referral hospital with multiple specialties. Major rare disease areas include ectodermal dysplasias, skeletal dysplasias, neurological disorders, lysosomal storage disorders, and coagulation disorders, among others.
17	Kalawati Saran Children's Hospital, New Delhi	Pediatric disorders
18	King Edward Memorial (KEM) Hospital, Mumbai	Hemoglobinopathies, LSDs, inborn errors of metabolism
19	LV Prasad Eye Institute, Hyderabad	Eye diseases
20	Manipal University, Manipal	A referral hospital with multiple specialties. Major rare disease areas include skeletal dysplasia, neurodegenerative diseases, metabolic disorders, bleeding disorders, and malformation syndromes, among others
21	Maulana Azad Medical College, New Delhi	LSDs, skeletal dysplasia, hemophilia, pediatric disorders
22	National Institute of Biomedical Genomics (NIBMG), Kalyani	Hemoglobinopathies, Wilson disease, DMD, eye disorders
23	National Institute of Mental Health and Neuro-Sciences (NIMHANS), Bangalore	Various rare neuromuscular disorders including limb girdle muscular dystrophy, amyotrophic lateral sclerosis, metabolic myopathies, rare congenital myasthenic syndromes, neuropsychiatric syndromes, SCA, mitochondrial disorders, and metabolic disorders
24	Nizam's Institute of Medical Sciences, Hyderabad	A referral hospital for a variety of rare genetic disease such as LSDs, hemoglobinopathies, and neurodegenerative disorders, among others
25	Osmania University, Hyderabad	Rare chromosomal disorders
26	Post Graduate Institute of Medical Education and Research (PGIMER), Chandigarh	A tertiary referral hospital with multiple specialties. Major rare disease areas include immune diseases, bone diseases, and Wilson diseases, among others
27	Sankara Nethralaya, Chennai	Eye diseases

**Table 3** List of major research centers working on rare diseases in India (Continued)

S. no.	Research centers	Major rare disease research areas
28	Safdarjung Hospital, New Delhi	Metabolic diseases, cardiac disorders
29	Sanjay Gandhi Postgraduate Institute of Medical Sciences (SGPGIMS), Lucknow	A referral hospital with multiple specialties. Major rare disease areas include LSDs, oro-facial-digital syndromes, neurodevelopment disorders, hemoglobinopathies, neurodegenerative disorders, and metabolic disorders, among others
30	Sir Ganga Ram Hospital, New Delhi	A referral hospital for LSDs, neurodegenerative disorders, IEMs, mitochondrial disorders, and other rare diseases.
31	Sri Chitra Tirunal Institute of Medical Sciences and Technologies, Thiruvananthapuram	A referral hospital for immune diseases, autoinflammatory diseases, LSDs, and cardiac diseases, among others
32	The Centre for DNA Fingerprinting and Diagnostics (CDFD), Hyderabad	Neuromuscular disorders, metabolic diseases, hemoglobinopathies, thrombotic disorders, triplet repeat disorders, and LSDs, among others
33	The Datta Meghe Institute of Medical Sciences (DMIMS), Wardha	Multiple rare diseases
34	All India Institute of Medical Sciences, Jodhpur	A referral hospital for several rare diseases including cystic fibrosis, and leucocyte adhesion defect, among others
35	University of Delhi South Campus, New Delhi	Inborn errors of metabolism, intellectual disability, and Parkinson's disease

**Table 4** A comprehensive list of rare disease organizations and resources that provide patient support [modified from [90]]

S. no.	Organizations	Websites
1	Alzheimers and Related Disorders Society Of India (ARDSI)	<a href="http://ardsi.org/">http://ardsi.org/</a>
2	Birth Defects Registry of India	<a href="http://www.fcrf.org.in/bdri_abus.asp">http://www.fcrf.org.in/bdri_abus.asp</a>
3	Down Syndrome Federation India	<a href="http://downsyndrome.in/">http://downsyndrome.in/</a>
4	Fragile X Society–India	<a href="http://www.fragilex.in/">http://www.fragilex.in/</a>
5	Genetic Alliance	<a href="http://www.geneticalliance.org">http://www.geneticalliance.org</a>
6	Hemophilia Federation	<a href="http://www.hemophilia.in/">http://www.hemophilia.in/</a>
7	Indian Rett Syndrome Foundation	<a href="http://www.rettssyndrome.in">www.rettssyndrome.in</a>
8	Indian Association of Muscular Dystrophy	<a href="http://www.iamd.in">www.iamd.in</a>
9	Indian Prader-Willi Syndrome Association	<a href="https://ipwsa.in/">https://ipwsa.in/</a>
10	Indian Patients Society for Primary Immunodeficiency (IPSPI)	<a href="http://www.ipspiindia.org">www.ipspiindia.org</a>
11	Indian Organization for Rare Diseases (I-ORD)	<a href="http://www.i-ord.org/">http://www.i-ord.org/</a>
12	Indian Society for Primary Immune Deficiency	<a href="http://www.ispid.org.in/">http://www.ispid.org.in/</a>
13	Lysosomal Storage Disorders Support Society (LSDSS)	<a href="http://www.lsdss.org">www.lsdss.org</a>
14	Metabolic Errors and Rare Diseases (MERD)	<a href="http://merdindia.com">http://merdindia.com</a>
15	Muscular Dystrophy Association India	<a href="http://mdindia.net/">http://mdindia.net/</a>
16	Muscular Dystrophy Foundation India	<a href="http://www.mdfindia.org">http://www.mdfindia.org</a>
17	Muskaan (intellectually disabled)	<a href="https://muskaanthengo.org/">https://muskaanthengo.org/</a>
18	National Thalassemia Welfare Society	<a href="http://www.thalassemiaindia.org/">http://www.thalassemiaindia.org/</a>
19	Open Platform for Rare Diseases (OPFORD)	<a href="https://opford.org/">https://opford.org/</a>
20	Organization of Rare Disorder India (ORDI)	<a href="https://ordindia.org/">https://ordindia.org/</a>
21	Pompe Foundation	<a href="http://pompeindia.org/">http://pompeindia.org/</a>
22	Rare Diseases India	<a href="http://www.rarediseasesindia.org">http://www.rarediseasesindia.org</a>
23	Retina India	<a href="http://retinaindia.blogspot.com/">http://retinaindia.blogspot.com/</a>
24	Sjogren's India	<a href="http://www.sjogrenindia.org">http://www.sjogrenindia.org</a>
25	Society for Hemophilia Care, India	<a href="http://www.shcindia.org/">http://www.shcindia.org/</a>
26	Thalassemics India	<a href="http://www.thalassemicsindia.org">www.thalassemicsindia.org</a>



reported variation(s). Genomics for Understanding Rare Diseases: India Alliance Network (GUARDIAN) is a research consortium which was proposed to address the above challenges. The consortium includes clinicians, clinical geneticists, genomics scientists, computational analysts, and basic research biologists, among others. The clinicians and clinical geneticists form the primary contacts and act as caregivers for the patients. The geneticists, genomics scientists, and researchers provide the necessary expertise required to identify the genetic variations, create models for understanding disease mechanisms, and explore the therapeutic potential of small molecules for rare genetic diseases. The simplified workflow of the GUARDIAN consortium is summarized in Fig. 1. The GUARDIAN is an open-ended consortium of individuals, who are actively invited to join the consortium, with an agreement to follow the general principles and framework, and the data access policies. A common framework for the exchange of datasets, resources within the consortium, and participatory approach has been proposed to realize the full potential of clinical genomics.

The aim of the GUARDIAN consortium is to establish a unique collaborative framework in health care planning, implementation, and delivery in the specific area of rare genetic diseases. The consortium proposes to apply the power of genomics for systematic characterization and diagnosis of rare genetic diseases in India. The GUARDIAN network is connected to hospitals and major tertiary care centers across India. The consortium currently encompasses over 240 clinicians/researchers, from 70 clinical/research centers across India [132]. The GUARDIAN is a research program and not a clinical service.

#### **GUARDIAN ethical framework**

A strong foundation of an ethical and legal framework is necessary for seamless collaboration and sharing of genetic data across the boundaries of institutions. The GUARDIAN consortium is strongly anchored on the basic principles of beneficence, reciprocity, justice, and professional responsibility. As part of the collaborators' network, a common format for collection of clinical and genetic data has been created. Additional efforts have



gone into standardizing the patient information. The benefits and potential ethical, legal, and social implications of whole exome or genome sequencing and availability of the anonymized data in the public domain are conveyed in detail to the patients and family. The identity stripped clinically annotated data of variations is available to all the members through a firewalled access. In addition, publications in peer-reviewed journals serve as the major interaction points for sharing findings with the general clinical and research community.

#### ***GUARDIAN clinical registry***

As part of the collaborative initiative, a referral system for systematic collection and curation of baseline data is being maintained. The program collects detailed clinical information, including the signs, symptoms, and clinical investigations performed on the patient and family members. The GUARDIAN maintains a semantically oriented framework, which relies extensively on the internationally accepted and popularly used semantic ontologies established and widely used including the human phenotype ontology [133]. The application of such a centralized data resource is manifold. While on the one end, it not only provides a holistic view of the burden of genetic diseases in the country, it also provides immense insights into the common and rare genetic variants in different sub-populations. This would enable clinicians and policy-makers to design intervention programs including genetic education and genetic counseling.

#### ***GUARDIAN sequence data generation***

A centralized sequencing facility has been established at the CSIR-Institute of Genomics and Integrative Biology (CSIR-IGIB), Delhi, which can be accessed by any collaborator in order to generate high-quality NGS sequencing data as per international standards [134–136], with various platforms such as Hiseq 2500 and NovaSeq 6000 (Illumina Inc. USA). A dedicated training team for both experimental and computational work necessary to perform the data capture and analysis of high-throughput sequencing data is also channelized as a part of the GUARDIAN consortium. Investigators are free to generate sequence data on their own or from other commercial facilities that adhere to international guidelines and GUARDIAN consortium standards. The sequencing requirements are updated and modified in accordance with the technological advancement and emerging international consensus.

#### ***GUARDIAN data analysis, integration, interpretation, and sharing***

GUARDIAN stands for providing scientifically sound and clinically actionable solutions. The genomes/exomes of

patients are analyzed through custom built in-house bio-informatic pipelines to identify the most accurate genetic variation that can explain a certain condition. Further, the pathogenicity of variants is predicted by the latest guidelines laid down by the American College of Medical Genetics and Genomics [136]. The GUARDIAN consortium relies heavily on datasets, tools, and resources developed across the whole world, including methods and tools developed as part of the OpenPGx consortium [137, 138]. The consortium depends on open source architectures, tools, and open access resources, to enable easy replication, scalability, and future implementation in independent clinical setups.

Data sharing also forms a major component of the program and collaboration. The anonymized clinically annotated data of variations is available to all members through a firewalled access. In addition, the summary data of each novel variant and/or allele frequencies would be available in the public domain without access restrictions. Credits for contributions are a major point to address in such a scalable collaborative network. All collaborating members of the network shall agree to adhere to basic principles of data veracity and ethical codes of conduct. The credit-sharing agreement forms the major framework of trust between participating members. This shall be in line with principles laid out for biomedical resource contributions [139].

#### ***GUARDIAN reporting, community screening, and disease modeling***

Once the GUARDIAN computational analysis identifies a pathogenic variation of clinical significance, it is subjected to validation by segregation analysis. After this, if the identified genetic variation is immediately actionable, the information is transferred to the clinician as a research report which will be used for patient counseling. This genetic information can further be used for making informed decisions by the family. Wherever required, the genetic variation information is utilized for potential community-level screening programs, thus building towards affordable diagnostic solutions.

In the case where novel pathogenic variations are identified, researchers at the GUARDIAN consortium replicate the disease in suitable models such as zebrafish and patient-derived iPSCs to gain the correlation between the disease phenotype and the identified variant. Genetic engineering to create disease models also provides the opportunity for discovery of novel therapeutics as well as to repurpose existing drugs for new indications in rare genetic diseases.

#### ***GUARDIAN success stories***

A large number of cases have been solved through the GUARDIAN program, and a subset of interesting

investigations have been published in peer-reviewed journals, which encompass diseases as diverse as epidermolysis bullosa [140–143], familial Mediterranean fever [144], lamellar ichthyosis [145], sporadic acrokeratosis verruciformis [146], rare syndromes of mineralocorticoid excess [147], severe combined immunodeficiency [148], X-linked agammaglobulinemia [149], hyper IgE syndrome [150], Dowling-Degos disease [151], and megalencephalic leukoencephalopathy [152], to list a few. Furthermore, GUARDIAN is actively investigating the genetic conundrum in Indian rare disease cohorts conforming to cardiology, neurology, dermatology, primary immunodeficiency, endocrinology, nephrology, mitochondrial disorders, and lysosomal storage disorders, among others.

Of the many success stories of GUARDIAN, the diagnosis of a rare mutation in *megalencephalic leukoencephalopathy with subcortical cysts 1 (MLC1)* gene in leukodystrophy was instrumental in community service in the form of affordable diagnostics. Six children from a consanguineous Muslim family belonging to the Nalband community from north India were presented with difficulty in balancing the head and inability to sit independently, with recurrent episodes of seizures. Based on the clinical characteristics, the provisional diagnosis of leukodystrophy was made; however, leukodystrophies are a class of disorders with the involvement of multiple genes. Whole exome sequencing revealed a homozygous variation in the *MLC1* gene, found to be segregated among all the affected members and was absent in all the unaffected members. Based on this, the diagnosis of megalencephalic leukoencephalopathy with subcortical cysts (MLC) was confirmed. MLC is a rare leukodystrophy characterized by macrocephaly, progressive motor dysfunction, recurrent episodes of seizures, and mental retardation. Further, three more families from the same community were found to be affected and carried the same variation, indicating a founder effect. As a follow up for this, an additional 83 members of the community were screened. Out of these, 24 were found to be the carriers and 9 were affected [152]. The Nalband community consists of over 5000 members scattered across north India as well as Pakistan. Like many other communities in India, consanguineous marriages are common in the Nalband community. In order to aid the entire community, a polymerase chain reaction (PCR)-based assay for the Nalband mutation in *MLC1* has been developed for carrier status determination and prenatal screening, at an affordable cost.

Another area where the GUARDIAN has made a significant contribution is in the rare diseases of the skin. Epidermolysis bullosa (EB), a skin-blistering disease, was once considered ultra-rare in the Indian population. Epidermolysis bullosa simplex (EBS) is the most common subtype of EB. The GUARDIAN team identified a novel variant in the *Keratin 5 (KRT5)* gene in a large

multigenerational family from northwestern India. The variant was shown to be segregated in nine affected members in the family but found absent in five unaffected members. The study reported the first causative mutation for EBS from India [140]. Whole exome sequencing has also enabled the detection of a novel homozygous nonsense variant in *Keratin 14 (KRT14)* gene in an autosomal recessive form of EB, in two siblings presented with generalized blistering of the skin and dystrophic nails. The same study identified a known homozygous stop gain variant in the same gene in a child with trauma-induced blistering all over the body [153]. In cases of junctional epidermolysis bullosa (JEB) and dystrophic epidermolysis bullosa (DEB), the phenotype and genotype spectrum of the disease was described for the first time from India through collaborative efforts of GUARDIAN. JEB was studied in a small cohort of six patients from four consanguineous families with a wide range of clinical variability, identifying variations in the genes *laminin subunit alpha 3 (LAMA3)*, *laminin subunit beta 3 (LAMB3)*, *collagen type XVII alpha 1 (COL17A1)* [142]. In the case of DEB, 18 patients from 17 unrelated families were studied and 20 distinct variations were found in *COL7A1* gene [143]. There have also been other reports which discovered novel variants that expanded the known mutation spectrum of EB [141, 154].

GUARDIAN has contributed to the identification of the pharmacogenetic variants in *dihydropyrimidine dehydrogenase (DPYD)* gene, which determines the metabolism of the commonly used anti-neoplastic drug 5-fluorouracil, in south-east Asian countries [155]. The consortium has also undertaken international initiatives to derive the pharmacogenomic landscape in Malays [156] and Qatari populations [157, 158], and to identify genetic variants of Arab, Middle East, and North African populations [159, 160]. GUARDIAN has also set up a systematic pipeline for next generation sequencing of the mitochondrial genome for clinical applications, called the mit-o-matic [80].

In the era of clinical genomics, it is imperative for clinicians to be well equipped with the basics of high-throughput data analysis so as to interpret the data concerning a certain disease. Keeping this in mind, the GUARDIAN consortium initiated an outreach program, where clinicians are trained in basics of NGS technologies and systematic computational analysis of sequencing data as a part of continuing medical education (CME) workshops. A handbook called 'Exome Sequence Analysis and Interpretation for Clinicians' has been prepared and made available for free download from Google Books [161]. Over 8000 soft copies of the book have been downloaded and over 800 print copies have been distributed to clinicians in meetings and CMEs (as of January 2019). More than 500 clinicians

have been trained across the country. The GUARDIAN outreach program is a small step towards providing health and economic benefits to families with rare genetic diseases.

### Impact of genomics in diagnosis of rare genetic diseases in India

It has been increasingly shown that the challenges of genetic and phenotypic heterogeneity which makes diagnosis of rare genetic diseases cumbersome could potentially be addressed by using next generation sequencing techniques, enabling the high-throughput identification and annotation of causal variants [126, 129, 162, 163]. In the present scenario, the rare diseases which require immediate attention in India are primary immunodeficiencies, hemoglobinopathies, muscular dystrophies, metabolic disorders, and neurological disorders, among others. The earlier section described the contributions made by a genomics-enabled nation-wide network, GUARDIAN. There have also been other individual genomics-based studies that have aided in addressing rare diseases.

In the case of Duchenne muscular dystrophy (DMD), a wide spectrum of mutations and frequencies have been shown in patients from different Indian sub-populations [164–166]. The dystrophin gene spans over 2000 kb at the DNA level, with pathogenic variations identified within introns as well. Traditional methods based on multiplex ligation-dependent probe amplification (MLPA) have been used to detect carrier status in DMD [167–170]. A recent study showed that NGS can be used in the diagnosis of muscular dystrophies in MLPA negative cases with a success rate of as high as 100% [171].

Lysosomal storage disorders (LSD), a class of more than 50 genetic diseases, are found to be of high burden in India [172]. The overlapping phenotypes and involvement of multiple genes in lysosomal disorders, and the need for intervention in the form of enzyme replacement therapy at the earliest, call for use of NGS approaches for faster diagnosis. In Niemann–Pick disease type C, an LSD with a wide clinical spectrum, a novel mutation was identified by whole exome sequencing in a proband of Asian origin, which was a deletion spanning two exons of *Niemann–Pick disease type C2 (NPC2)* gene [173].

An estimated one million Indians are affected by primary immunodeficiencies, a class comprising of hundreds of genetic disorders [174]. The utmost challenging facet of PIDs is under diagnosis, owing to the high incidence of infectious diseases in countries like India [175]. Whole exome sequencing approach has proved to be instrumental in identifying mutations in capillary sequencing negative cases of X-linked agammaglobulinemia (XLA) [149], severe combined immunodeficiency (SCID) [148], B cell expansion with NF- $\kappa$ B, and T cell anergy

(BENTA) [176], apart from targeted next generation sequencing in SCID [177] and major histocompatibility complex class II deficiency [178].

Mitochondrial disorders are difficult to diagnose owing to overlapping phenotypes and multi-system involvement. Whole mitochondrial genome sequencing coupled with nuclear gene sequencing has been performed to establish genotype-phenotype correlations in a cohort of patients from South India [179]. Whole exome sequencing has incidentally helped in diagnosing mitochondrial diseases due to nuclear genome variations [180, 181].

In case of autosomal recessive forms of ataxia, such as spastic ataxia [182] and cerebellar ataxias [183], homozygosity mapping as well as whole exome sequencing has played a major role in discovering the novel variants in Indian patients. Application of genomic diagnosis has been appreciated for skeletal dysplasias in a recent study. The study on a large cohort using capillary sequencing as well as NGS has added novel variants to the existing literature [184]. Exome sequencing also has been used to discover novel mutations in multiple joint dislocation syndrome [185], Schwartz-Jampel syndrome type 1 [186], and progressive pseudorheumatoid dysplasia [187]. Currently, a limited number of clinicians are using NGS-based diagnosis of rare genetic diseases in India but this number is increasing at a rapid pace. With several success stories emerging from India, genomics will become a mainstay for diagnosis of rare genetic diseases in the near future.

### Translating genomics to affordable diagnostics for rare genetic diseases

Although the cost of next generation sequencing-based diagnostics is declining, with more than 70 million people suffering from a genetic disease in India, affordable and faster measures are required to cater to the needs of the ailing population. CSIR-IGIB has an ongoing outreach platform to provide affordable access to genetic testing for common genetic diseases. The program named “Genomics and other Omics tools for Enabling Medical Decision (GOMED)” [188] provides molecular genetic assays for clinical diagnosis, prenatal testing, and carrier screening. In this ‘from bench to bedside’ model, a battery of low-cost genetic diagnostic assays for diseases pertaining to neurology, cardiology, and many other disorders are available. Till now, over 90 candidate gene tests and 7 comprehensive gene panel tests have been developed by GOMED. Over 20,000 molecular tests for about 6000 patients have been performed across the country (As of 2018). This clinical service is provided free of cost to needy patients. GOMED has been particularly beneficial in the community screening of sub-population-specific mutations.

Whole exome sequencing had revealed a founder mutation in *MLC1* gene in individuals from Nalband community suffering from megalencephalic leukoencephalopathy with subcortical cysts (MLC) [152]. As part of GOMED, a low-cost diagnostic assay was developed to screen for carriers in other members of this community comprising of 5000 people scattered across different regions in north India. Spinocerebellar ataxia (SCA) type 3, known as Machado–Joseph disease (MJD) is one of the most common ataxias globally, while presenting rarely in India. Intervention by CSIR-IGIB revealed the hidden burden of SCA3/MJD in 100–200 families in a close-knit community in Maharashtra. This information is now available as an assay under GOMED. GOMED also expands to pharmacogenetic testing to prevent adverse reactions to commonly used drugs such as the anticancer drug 5-fluorouracil. 5-fluorouracil (5-FU) is an anti-neoplastic drug which is administered in a number of cancers, the clearance of which is mediated by a rate-limiting enzyme dihydropyrimidine dehydrogenase (DPYD). Genotyping of four variants in *DPYD* gene that were found to be associated with 5-FU toxicity in South Asian population [155] has been made available as an affordable diagnostic assay for testing cancer patients before administering the drug to prevent adverse reactions. The GOMED program also actively works with commercial diagnostic companies to provide technologies for the affordable diagnosis of common and rare genetic diseases in India.

As a step towards improving public health, efforts have also been undertaken to compile a directory of genetic test services and counseling centers in India. The directory includes about 120 centers across various states in India. It acts as a resource for clinicians as well as researchers for referring to facilities which provide accessible and comprehensive public healthcare [189].

### The way ahead

There are a few priority areas that are emerging in the country as far as rare diseases are concerned. Newborn screening at a nation-wide level is pivotal in reducing the burden of rare diseases. In 2014, India Newborn Action Plan (INAP) was released to reduce the incidence of child birth defects and stillbirths [190]. While at present, there are limitations in implementing genomics-based diagnosis at population scale [191], Indian pediatricians are hopeful about the genomic interventions and resultant advancements in diagnosis, especially for non-invasive prenatal testing [192]. National Policy for Treatment of Rare Diseases was released by the Indian Ministry of Health and Family Welfare in 2017 [193]. However, this policy was withdrawn in November 2018 to the utter dismay of the patients and family members suffering from rare diseases

[194]. As personal genome-sequencing becomes popular, it is important to create a policy and a legal framework for non-discrimination of individuals based on the genetic information. This would be in line with the Genetic Information Nondiscrimination Act (GINA) of the USA but also adapted to the social and cultural sensibilities specific to India. As we look ahead, we should involve stakeholders such as government policy-makers, research scientists, clinicians, hospitals, patient groups, and non-governmental organizations to join forces to find meaningful solutions for rare diseases patients.

For a large and heterogeneous population like that of India, it has been shown that the international genomics initiatives such as the 1000 genome project have an inadequate representation of the genetic diversity due to limited sampling [20]. In highly endogamous populations such as the Ashkenazi Jewish population, genomics has been crucial in understanding rare diseases with founder effects [195]. With an enormous and stratified population, practicing extensive endogamy [39], it is expected that India would have a high prevalence of rare genetic diseases. Therefore, it is essential to know the causal genes and pathogenic genetic variants and the sub-populations where they are prevalent, to aid in the appropriate and cost-effective diagnosis of rare diseases. There are several initiatives in India that are attempting to address this space by building large-scale whole genome datasets of the representative population. Programs such as the GenomeAsia100K, which has representative samples from India, seek to sequence and analyze individuals to help enable medical applications [196]. The Government of India has announced a Bioscience Mission for Precision Health and Optimal Well-being, which will involve large-scale human genome sequencing across India [197]. Towards this, the Council of Scientific and Industrial Research (CSIR), India, has also initiated a whole genome sequencing program titled “Genomics for Public Health (IndiGen)” [198] to help accelerate biomedical applications in India. These population scale genomics programs will definitely provide the momentum and ecosystem for driving rare disease genomics in India.

### Conclusion

India is home to culturally and genetically diverse populations, which are burdened by genetic diseases. Due to the high prevalence of recessive alleles owing to endogamous practices, rare diseases form a significant burden in India. Genomics can greatly aid in addressing rare disease burden by faster and more accurate diagnoses. The Genomics for Understanding Rare Diseases: India Alliance Network (GUARDIAN) provides a template for a nation-wide collaborative platform that uses the power

of genomics to dissect the rare disease conundrum. More such pan-India genomics-driven initiatives can help in deriving Indian-specific references for deducing pathogenic and benign variations in the population, which can pave the way for precision medicine, including in the rare disease space.

#### Abbreviations

5-FU: 5-Fluorouracil; AHF: Anti-hemophilic factors; BENTA: B cell expansion with NF- $\kappa$ B and T cell anergy; CFC: Clotting factor concentrates; CME: Continuing medical education; COL17A1: Collagen type XVII  $\alpha$ 1; CSIR: Council of Scientific and Industrial Research; DEB: Dystrophic Epidermolysis Bullosa; DMD: Duchenne Muscular Dystrophy; DPYD: Dihydropyrimidine dehydrogenase; EB: Epidermolysis bullosa; EBS: Epidermolysis bullosa simplex; GOMED: Genomics and other Omics tools for Enabling Medical Decision; GUARDIAN: Genomics for Understanding Rare Diseases; India Alliance Network; ICMR: Indian Council of Medical Research; IGDD: Indian Genetic Disease Database; IGB: Institute of Genomics and Integrative Biology; IGV: Indian Genome Variation; INAP: India Newborn Action Plan; INCAP: India Charitable Access Program; INDEX-db: INdian EXome database; JEB: Junctional Epidermolysis Bullosa; KRT: Keratin; LAMA3: Laminin subunit  $\alpha$ 3; LAMB3: Laminin subunit  $\beta$ 3; LSD: Lysosomal storage disorders; MDCRC: Molecular Diagnostics, Counseling, Care and Research Centre; MJD: Machado Joseph Disease; MLC: Megalencephalic Leukoencephalopathy with subcortical Cysts; MLPA: Multiplex ligation-dependent probe amplification; NGS: Next generation sequencing; NPC2: Niemann-Pick disease type C2; ORDI: Organization for Rare Diseases India; PCR: Polymerase chain reaction; SAGE: South Asian Genomes and Exomes; SCA: Spinocerebellar ataxia; SCID: Severe combined immunodeficiency; SMA: Spinal muscular atrophy; SNP: Single nucleotide polymorphism; XLA: X-linked agammaglobulinemia

#### Acknowledgements

The GUARDIAN Consortium is very grateful to all the patients, family members, and patient support groups for their active participation and cooperation during the course of the research study. The authors also thank the hospitals, clinical and research fraternity for joining hands together to solve the rare disease mysteries in India. SS and VS thank CSIR and IGB for constant support and encouragement. SS and VS thank M/S Sanofi Genzyme Pvt Ltd. for the unrestricted educational grant that permitted training of medical professionals in the area of tclinical genomics.

#### The GUARDIAN Consortium:

Anjali Bajaj<sup>5</sup>, Samatha Mathew<sup>5</sup>, Shamsudheen Karuthedath Vellarikkal, Ambily Sivasdas, Rahul C Bhojar, Kandarp Joshi, Abhinav Jain, Anushree Mishra, Ankit Verma, Rijith Jayarajan, A Nalini, A Ravi Kumar, A.T Arasar Seeralar, Aayush Gupta, Achal K Srivastava, Aditi Joshi, Aditi Sinha, Aditya Jandial, Afreen Khan, Akhilesh K Sonakar, Alex Chandy, Aman Sharma, Ambuj Roy, Amit Rawat, Amitabh Biswas, Andrew Vanlalawma, Anita Chaudhary, Anita Chopra, Ankit Panday, Ankit Sabharwal, Ankita Mitra, Ankita Narang, Anna Rajab, Anoop Kumar, Anoop Singh Gurjar, Anop Singh Ranawat, Anu R I, Anup Kumar Tiwary, Anuradha, Aquil Kalanad, Aradhana Mathur, Arjun Lakshman, Arushi Batra, Arvind Bagga, Ashish Aggarwal, Ashok Gupta, Ashu Rastogi, Aslam PK, Astha V, Aswin Nair, Athulya E P, Atri Chatterjee, Atul Jindal, Atul Kumar Kashyap, B Priyadarshini, Babu Ram Thapa, Balram Bhargava, Balram Sharma, Bani Jolly, Bharath Ram Uppilli, Bharathi Balachander, Bhim Shankar, Bibhas Kar, Binukumar B K, C. Lalchandama, Chaitanya Datar, Chetana Sachidanandan, D C Master, Daisy Khera, Debashish Chowdhury, Debashish Danda, Deepak Kumar, Deepika Pandhi, Deepthi Siddharthan, Disha Sharma, Divya Pachat, Brijesh Sharma, Durga Rao Vegulada, GRSRNN Naidu, G Padma, G.Vishnu Priya, Gautam Sharma, Gauthamen R, Geeta Govindaraj, George M Varghese, Gireesh S, Gopikrishnan Unnikrishnan, Hafiz SA, Hazeena KR, Heena Dhiman, Hema Singh, Hrishikesh Sarkar, Istaq Ahmed, Jagadeesh Menon, Jatinder Goraya, Jennifer Mathew, Jineesh Thottath, Jitendra K Sahu, Jitendra Oswal, John Menachery, Judith Mary Hariprakash, K Bhargava, K K Talwar, K M Cherian, K P Aravindan, K Pramila, K Saroja, K Shantaraman, Kavita Pandhare, Kiran Kumar Mandapati, Kiran P, Kotha Rakesh, Krati Shah, Krishnan C, Kriti Shah, Kuldeep Singh, Kuljeet Anand, Lalawmpuii Pachuau, Laxmisha Chandrashekar, Liza Rajasekhar, Lopamudra Mishra, M V Padma, Madhulika Kabra, Madhumita Roy Chowdhary, Malika Seth, Maneesh Rai, Manish Kumar, Manish Parakh,

Manisha Goyal, Manisha Gurjar, Manisha Sahay, Mercy Rophina, Mitali Mukerji, Mohammed Ali, Mohammed Faruq, Mohandas Nair Karipphoth, Mohit Kumar Divakar, MP Jayakrishnan, Mukesh Kumar, Mukta Poojary, Mukund A Prabhu, Nachimuthu Senthil Kumar, Nadeem Rais, Nalini Bhaskaranand, Narendra Kumar Bagri, Naveen Sankhyan, Neeraj Awasthy, Neeraj Gupta, Neeraj Parakh, Neeraja Gupta, Neetu Bhari, Neetu Kushwaha, Neha Sharma, Neha Virmani, Nilanjan Kundu, Nishad Plakkal, Nishu Tyagi, Nita Radhakrishnan, Nitish Naik, Nitish Rai, Nivedita Mondal, Nupur Bhargava, Pankaj Hari, Paras Sehgal, Piyush Kumar, Pooja Chauhan, Pooja Mailankody, Pooja Sharma, Poonam Parakh, Pragya A Nair, Praloy Chakraborty, Prasanna Kumar Shirol, Pratibha Singh, Pratosh Gangadhar, Prawin Kumar, Purna chandra, R Krishnan, R Srilakshmi, R Sriranga Lakshmi, R. Anantharaman, Radha Mahadevan, Rahul Mahajan, Rajasubramaniam Shanmugam, Rajat Sharma, Rajendran V R, Rajinder K Dhamija, Rajit Pillai Ramanan, Rajive Kumar, Rajneesh A R, Rajnish Juneja, Rakesh Aggarwal, Rakesh Sahay, Ramakrishnan S, Ranjith Narayanan, Ravindra Shukla, Remya Koshy, Renu Kumari, Richa Chaudhary, Richa Jain, Riyaz Arakkal, Roopa Rajan, Rowmika Ravi, S Baruah, S. Sitaraman, Sadandandavalli Retnaswami Chandra, Saia Chenkual, Sailaja V, Sakshi Ambawat, Samhita Panda, Sana Zahra, Sanchit Kumar, Sandeep Arora, Sandeep Mathur, Sandeep Seth, Sandhya P, Sangam Goswami, Sangita Paul, Sanjay Pandey, Santharaman Kalyanaraman, Saroj Patnaik, Saruchi Wadhwa, Sathi Venu, Satyan Nanda, Saumya Panda, Saurabh Chopra, Saurabh Singh, Savinitha P, Seema Kapoor, Sesh Sivadasan, Sethuraman G, Shaista Parveen Khan, Shaji CV, Shanmugam Gurusamy, Sheffali Gulati, Shrey Gandhi, Sivaprakash Ramalingam, Smita Nath, Suresh Kumar, Sona Sathian, Sonal Lakhani, Soumya S Nair, Soumya Sundaram, Sourav Ghosh, Sree Bhushan Raju, Sreejith Valappil, Sreelata Nair, Srikanth Kadyada Puttaiah, Sruthi S Nair, Suja K Geevarghese, Sujata Mohanty, Sujay Khandpur, Suman Jain, Sumeet, Sumit Sharma, Suruchi Trehan, Suvasini Sharma, Sweta Jain, Swetha Jain, Tarun Kumar Badam, Umamaheswari S, Utkarsh Gaharwar, Uzma Shamim, Vadlamudi Raghavendra Rao, Vamsi Krishna, Vandana Jain, Varun Suroliya, Varuna Vyas, Veena Vedartham, Venketesh S, Vigneshwar Senthivel, Vijaykumar Bhavi, Vilas Jadhav, Vinay Gera, Vishal Dixit, Vishal Gupta, Vishnu Agarwal, Vishnu V Y, Vishu Gupta, Vysakha K V, Yugal K Sharma, Samir K Brahmachari, Vinod Scaria\*, Sridhar Sivasubbu\*. <sup>5</sup>These authors contributed equally.

#### Authors' contributions

All authors read and approved the final manuscript.

#### Funding

The rare disease program at CSIR-IGIB is funded through the following research grants from the Council of Scientific and Industrial Research (CSIR), India 2012–2016: grant no. BSC0212 and BSC0122; September 2016 onwards: grant no. MLP1601; July 2018 onwards: grant nos. MLP1801, MLP1802, and MLP1809.

#### Availability of data and materials

Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

#### Ethics approval and consent to participate

Not applicable.

#### Consent for publication

Not applicable.

#### Competing interests

- I. Sridhar Sivasubbu and Vinod Scaria have received an unrestricted educational grant from M/S Sanofi Genzyme Pvt Ltd. for training medical professionals in the area of clinical genomics.
- II. Sridhar Sivasubbu and Vinod Scaria collaborate with M/S Adams Genetics Pvt Ltd. and M/S Genique Lifesciences Pvt Ltd. for developing technologies for interpretation of genome-scale datasets.
- III. Sridhar Sivasubbu, Vinod Scaria, and Md Faruq have developed and transferred molecular assays for diagnosis of genetic diseases to M/S Lal Path Labs Pvt Ltd.
- IV. Sridhar Sivasubbu and Vinod Scaria have developed and transferred NGS-based assays and computational reporting engine for

diagnosis of mitochondrial diseases to M/S Eurofins Clinical Genetics India Pvt Ltd.

#### Author details

<sup>1</sup>CSIR Institute of Genomics and Integrative Biology, Delhi 110025, India. <sup>2</sup>Academy of Scientific and Innovative Research, Delhi, India. <sup>3</sup>National Institute of Mental Health and Neuro Sciences, Bengaluru, India. <sup>4</sup>Thalassemia and Sickle Cell Society, Hyderabad, India. <sup>5</sup>Institute of Child Health and Hospital for Children, Chennai, India. <sup>6</sup>Dr. DY Patil Medical College and hospital, Pune, India. <sup>7</sup>All India Institute of Medical Sciences, Delhi, India. <sup>8</sup>Post Graduate Institute of Medical Education and Research, Chandigarh, India. <sup>9</sup>Govt Medical college Kozhikode, Kerala, India. <sup>10</sup>Galgotias University, Greater Noida, India. <sup>11</sup>Mizoram University, Aizawl, India. <sup>12</sup>Sawai Man Singh Medical College, Jaipur, India. <sup>13</sup>Wadia Children Hospital, Mumbai, India. <sup>14</sup>VPS Healthcare, Abu Dhabi, United Arab Emirates. <sup>15</sup>Government Medical College, Barmer, India. <sup>16</sup>MVR Cancer Centre and Research Institute, Calicut, India. <sup>17</sup>Government Medical College, Haldwani, India. <sup>18</sup>Government Medical College Kozhikode, Kozhikode, India. <sup>19</sup>JK Lone Hospital, Sawai Man Singh Medical College, Jaipur, India. <sup>20</sup>Christian Medical College Hospital, Vellore, India. <sup>21</sup>Vardhman Mahavir Medical College & Safdurjung Hospital, Delhi, India. <sup>22</sup>All India Institute of Medical Sciences, Raipur, India. <sup>23</sup>St. Johns Medical College Hospital, Bangalore, India. <sup>24</sup>Lakeshore Hospital, Kochi, India. <sup>25</sup>Madras Medical Mission, Chennai, India. <sup>26</sup>Department of Pathology, Civil Hospital Aizawl, Aizawl, India. <sup>27</sup>Sahyadri Medical Genetics and Tissue Engineering Facility, Pune, India. <sup>28</sup>Medical College, Baroda, India. <sup>29</sup>All India Institute of Medical Sciences, Jodhpur, India. <sup>30</sup>GB Pant Hospital, Delhi, India. <sup>31</sup>University College of Medical Sciences and Guru Teg Bahadur Hospital, Delhi, India. <sup>32</sup>Malabar Institute of Medical Sciences, Calicut, India. <sup>33</sup>Dr. Ram Manohar Lohia hospital, Delhi, India. <sup>34</sup>Genes N life Healthcare, Hyderabad, India. <sup>35</sup>Christian Medical College, Vellore, India. <sup>36</sup>sree Chithra Tirunal Institute of Medical Sciences and Technology, Trivandrum, India. <sup>37</sup>General Hospital, Thiruvananthapuram, India. <sup>38</sup>Noorul Islam Multi Speciality Hospital, Thiruvananthapuram, India. <sup>39</sup>Sawai Man Singh Hospital, Jaipur, India. <sup>40</sup>Holy Family Hospital, Mumbai, India. <sup>41</sup>Dayanand Medical College & Hospital, Ludhiana, India. <sup>42</sup>Calicut Medical College, Calicut, India. <sup>43</sup>Government Medical College Kozhikode, Calicut, India. <sup>44</sup>Bharati Vidyapeeth Medical College and Hospital, Pune, India. <sup>45</sup>Rajagiri Hospital, Aluva, India. <sup>46</sup>Max Super Speciality Hospital, Delhi, India. <sup>47</sup>Frontier Lifeline and Dr. K. M. Cherian Heart Foundation, Chennai, India. <sup>48</sup>Tirunelveli Medical College, Tirunelveli, India. <sup>49</sup>Malabar Cancer Centre, Kannur, India. <sup>50</sup>Osmania Medical College, Hyderabad, India. <sup>51</sup>ONE Centre for Rheumatology & Genetics, Vadodara, India. <sup>52</sup>Institute of Maternal & Child Health, Government Medical College, Kozhikode, India. <sup>53</sup>Isha Hospital, Vadodara, India. <sup>54</sup>Civil Hospital Aizawl, Aizawl, India. <sup>55</sup>Jawaharlal Institute of Postgraduate Medical Education and Research, Puducherry, India. <sup>56</sup>Nizam's Institute of Medical Sciences, Hyderabad, India. <sup>57</sup>Kasturba Medical College, Mangalore, India. <sup>58</sup>Dr. Sampurnanand Medical College, Jodhpur, India. <sup>59</sup>J K Lone Hospital, Sawai Man Singh Hospital, Jaipur, India. <sup>60</sup>Amrita institute of medical sciences, Cochin, India. <sup>61</sup>Chowpaty Medical Center, Mumbai, India. <sup>62</sup>Hitech Medicare Hospital, Udipi, India. <sup>63</sup>Max Hospital, Delhi, India. <sup>64</sup>Lady Hardinge Medical College, Delhi, India. <sup>65</sup>Sakhiya Skin Clinic, Surat, India. <sup>66</sup>Dr. Nilanjan Kundu's Nursing Home, Bankura, India. <sup>67</sup>Jawaharlal Institute of Postgraduate Medical Education & Research, Puducherry, India. <sup>68</sup>Super Speciality Pediatric Hospital and PG Teaching Institute, Noida, India. <sup>69</sup>Katihar Medical College, Dalan, India. <sup>70</sup>National Institute of Mental Health and Neurosciences, Bengaluru, India. <sup>71</sup>Pramukhswami Medical College, Karamsad, India. <sup>72</sup>Safdarjung Hospital, Delhi, India. <sup>73</sup>Organization for Rare Diseases, Jodhpur, India. <sup>74</sup>PVS Hospital, Calicut, India. <sup>75</sup>MES Medical College, Malappuram, India. <sup>76</sup>The Tamil Nadu Dr. M.G.R. Medical University, Chennai, India. <sup>77</sup>Frontier Lifeline Hospital, Chennai, India. <sup>78</sup>ICMR-National Institute of Research in Tribal Health, Jabalpur, India. <sup>79</sup>Base Hospital Delhi Cantt, Delhi, India. <sup>80</sup>Aster Malabar Institute of Medical Sciences Hospital, Kozhikode, India. <sup>81</sup>Osmania General Hospital, Hyderabad, India. <sup>82</sup>KMCT Medical College Hospital, Kozhikode, India. <sup>83</sup>Kannur Medical College, Kannur, India. <sup>84</sup>Tezpur University, Tezpur, India. <sup>85</sup>Maulana Azad Medical College, Delhi, India. <sup>86</sup>Command Hospital Air Force, Bengaluru, India. <sup>87</sup>Govind Ballabh Pant Institute of Postgraduate medical education and research, Delhi, India. <sup>88</sup>Army Research and Referral Hospital, Delhi, India. <sup>89</sup>KPC Medical College, Kolkata, India. <sup>90</sup>BLK Super Speciality Hospital, Delhi, India. <sup>91</sup>Maulana Azad Medical College, Lok Nayak Jai Prakash Narayan Hospital, Delhi, India. <sup>92</sup>TD Medical College, Allappuzha, India. <sup>93</sup>Kannur Dental College, Kannur, India. <sup>94</sup>Lifeline Super

Specialty Hospital, Pathanamthitta, India. <sup>95</sup>Mar Baselios Medical Mission Hospital, Ernakulam, India. <sup>96</sup>Lady Harding Medical College, Delhi, India. <sup>97</sup>Devadoss Multi Speciality Hospital, Madurai, India. <sup>98</sup>Pudhuvai Life Line Hospital, Puducherry, India. <sup>99</sup>Osmania University, Hyderabad, India. <sup>100</sup>Sri Satya Sai Institute of Higher Learning, Puttaparthi, India. <sup>101</sup>Aarya Child Epilepsy and Neurology Clinic, Kolhapur, India. <sup>102</sup>Base Hospital, Lucknow, India. <sup>103</sup>Greechithra Tirunal Institute of Medical Sciences and Technology, Trivandrum, India.

Received: 3 April 2019 Accepted: 26 June 2019

Published online: 25 September 2019

#### References

- Reich D, Thangaraj K, Patterson N, Price AL, Singh L. Reconstructing Indian population history. *Nature*. 2009;461(7263):489–94.
- Majumder PP, Basu A. A genomic view of the peopling and population structure of India. *Cold Spring Harb Perspect Biol*. 2014;7(4):a008540.
- Basu A, Sarkar-Roy N, Majumder PP. Genomic reconstruction of the history of extant populations of India reveals five distinct ancestral components and a complex structure. *Proc Natl Acad Sci U S A*. 2016;113(6):1594–9.
- Narasimhan AVM, Patterson N, Moorjani P, Lazaridis I. The Genomic Formation of South and Central Asia. 2018. <https://doi.org/10.1101/292581>.
- Singh KS. *People of India: Introduction*. Delhi: Oxford University Press; 2002.
- Bamshad M, Kivisild T, Watkins WS, Dixon ME, Ricker CE, Rao BB, et al. Genetic evidence on the origins of Indian caste populations. *Genome Res*. 2001;11(6):994–1004.
- Malhotra KC. Morphological composition of the people of India. *J Hum Evol*. 1978;7(1):45–53.
- Kivisild T, Rootsi S, Metspalu M, Mastana S, Kaldma K, Parik J, et al. The genetic heritage of the earliest settlers persists both in Indian tribal and caste populations. *Am J Hum Genet*. 2003;72(2):313–32.
- Cavalli-Sforza LL, Menozzi P, Piazza A. *The History and Geography of Human Genes: Abridged paperback Edition*. Princeton: Princeton University Press; 2018.
- Indian Genome Variation Consortium. The Indian Genome Variation database (IGVdb): a project overview. *Hum Genet*. 2005;118(1):1–11.
- Das K, Malhotra KC, Mukherjee BN, Walter H, Majumder PP, Papiha SS. Population structure and genetic differentiation among 16 tribal populations of Central India. *Hum Biol*. 1996;68(5):679–705.
- Majumder PP, Roy B, Banerjee S, Chakraborty M, Dey B, Mukherjee N, et al. Human-specific insertion/deletion polymorphisms in Indian populations and their possible evolutionary implications. *Eur J Hum Genet*. 1999;7(4):435–46.
- Thanseem I, Thangaraj K, Chaubey G, Singh VK, Bhaskar LVKS, Reddy BM, et al. Genetic affinities among the lower castes and tribal groups of India: inference from Y chromosome and mitochondrial DNA. *BMC Genet*. 2006;7:42.
- Khan F, Pandey AK, Borkar M, Tripathi M, Talwar S, Bisen PS, et al. Effect of sociocultural cleavage on genetic differentiation: a study from North India. *Hum Biol*. 2008;80(3):271–86.
- Borkar M, Ahmad F, Khan F, Agrawal S. Paleolithic spread of Y-chromosomal lineage of tribes in eastern and northeastern India. *Ann Hum Biol*. 2011;38(6):736–46.
- Guha P, Das A, Dutta S, Bhattacharjee S, Chaudhuri TK. Study of genetic diversity of KIR and TLR in the Rabhas, an endogamous primitive tribe of India. *Hum Immunol*. 2015;76(11):789–94.
- Chinniah R, Vijayan M, Thirunavukkarasu M, Mani D, Raju K, Ravi PM, et al. Polymorphic Alu insertion/deletion in different caste and tribal populations from South India. *PLoS One*. 2016;11(6):e0157468.
- Vadivelu MK. Emergence of sociocultural norms restricting intermarriage in large social strata (endogamy) coincides with foreign invasions of India. *Proc Natl Acad Sci USA*. 2016;113(16):E2215–7.
- Bhasin MK, Nag S. Consanguinity and its effects on fertility, mortality and morbidity in the Indian Region: A Review. *Int J Hum Genet*. 2012;12(4):197–301.
- Sengupta D, Choudhury A, Basu A, Ramsay M. Population stratification and underrepresentation of Indian subcontinent genetic diversity in the 1000 genomes project dataset. *Genome Biol Evol*. 2016;8(11):3460–70.
- Indian Genome Variation Consortium. Genetic landscape of the people of India: a canvas for disease gene exploration. *J Genet*. 2008;87(1):3–20.
- Narang A, Das RR, Chaurasia A, Mukhopadhyay A, Mukerji M, Dash D. IGVBrowser—a genomic variation resource from diverse Indian populations. *Database (Oxford)*. 2010;2010:baq022.

23. Sinha S, Mishra SK, Sharma S, Patibandla PK, Mallick PK, Sharma SK, et al. Polymorphisms of TNF-enhancer and gene for FcγRIIIa correlate with the severity of falciparum malaria in the ethnically diverse Indian population. *Malar J*. 2008;7:13.
24. Sinha S, Qidwai T, Kanchan K, Anand P, Jha GN, Pati SS, et al. Variations in host genes encoding adhesion molecules and susceptibility to falciparum malaria in India. *Malar J*. 2008;7:250.
25. Sinha S, Arya V, Agarwal S, Habib S. Genetic differentiation of populations residing in areas of high malaria endemicity in India. *J Genet*. 2009;88(1):77–80.
26. Jha P, Sinha S, Kanchan K, Qidwai T, Narang A, Singh PK, et al. Deletion of the APOBEC3B gene strongly impacts susceptibility to falciparum malaria. *Infect Genet Evol*. 2012;12(1):142–8.
27. Kanchan K, Pati SS, Mohanty S, Mishra SK, Sharma SK, Awasthi S, et al. Polymorphisms in host genes encoding NOSII, C-reactive protein, and adhesion molecules thrombospondin and E-selectin are risk factors for Plasmodium falciparum malaria in India. *Eur J Clin Microbiol Infect Dis*. 2015; 34(10):2029–39.
28. Bhattacharjee A, Banerjee D, Mookherjee S, Acharya M, Banerjee A, Ray A, et al. Leu432Val polymorphism in CYP1B1 as a susceptible factor towards predisposition to primary open-angle glaucoma. *Mol Vis*. 2008;14:841–50.
29. Kumar J, Garg G, Kumar A, Sundaramoorthy E, Sanapala KR, Ghosh S, et al. Single nucleotide polymorphisms in homocysteine metabolism pathway genes: association of CHDH A119C and MTHFR C677T with hyperhomocysteinemia. *Circ Cardiovasc Genet*. 2009;2(6):599–606.
30. Aggarwal S, Negi S, Jha P, Singh PK, Stobdan T, Pasha MAQ, et al. EGLN1 involvement in high-altitude adaptation revealed through genetic analysis of extreme constitution types defined in Ayurveda. *Proc Natl Acad Sci U S A*. 2010;107(44):18961–6.
31. Aggarwal S, Gheware A, Agrawal A, Ghosh S, Prasher B, Mukerji M. Combined genetic effects of EGLN1 and VWF modulate thrombotic outcome in hypoxia revealed by Ayurgenomics approach. *J Transl Med*. 2015;13:184.
32. Biswas A, Maulik M, Das SK, Ray K, Ray J. Parkin polymorphisms: risk for Parkinson's disease in Indian population. *Clinical genetics*. Denmark. 2007;72:484–6.
33. Biswas A, Sadhukhan T, Majumdar S, Misra AK, Das SK, Variation Consortium IG, et al. Evaluation of PINK1 variants in Indian Parkinson's disease patients. *Parkinsonism Relat Disord*. 2010;16(3):167–71.
34. Gupta A, Maulik M, Nasipuri P, Chattopadhyay I, Das SK, Gangopadhyay PK, et al. Molecular diagnosis of Wilson disease using prevalent mutations and informative single-nucleotide polymorphism markers. *Clin Chem*. 2007;53(9): 1601–8.
35. Chaki M, Sengupta M, Mondal M, Bhattacharya A, Mallick S, Bhadra R, et al. Molecular and functional studies of tyrosinase variants among Indian oculocutaneous albinism type 1 patients. *J Invest Dermatol*. 2011;131:260–2.
36. Grover S, Gourie-Devi M, Baghel R, Sharma S, Bala K, Gupta M, et al. Genetic profile of patients with epilepsy on first-line antiepileptic drugs and potential directions for personalized treatment. *Pharmacogenomics*. 2010; 11(7):927–41.
37. Talwar P, Kanojia N, Mahendru S, Baghel R, Grover S, Arora G, et al. Genetic contribution of CYP1A1 variant on treatment outcome in epilepsy patients: a functional and interethnic perspective. *Pharmacogenomics J*. 2017;17(3): 242–51.
38. Giri AK, Khan NM, Grover S, Kaur I, Basu A, Tandon N, et al. Genetic epidemiology of pharmacogenetic variations in CYP2C9, CYP4F2 and VKORC1 genes associated with warfarin dosage in the Indian population. *Pharmacogenomics*. 2014;15(10):1337–54.
39. Nakatsuka N, Moorjani P, Rai N, Sarkar B, Tandon A, Patterson N, et al. The promise of discovering population-specific disease-associated genes in South Asia. *Nat Genet*. 2017;49(9):1403–7.
40. Christianson A, Howson CP, Modell B. March Of Dimes Global Report On Birth Defects. 2006. Available from: <https://www.marchofdimes.org/global-report-on-birth-defects-the-hidden-toll-of-dying-and-disabled-children-full-report.pdf>. Accessed 28 Mar 2019.
41. Kaur A, Singh JR. Chromosomal abnormalities: genetic disease burden in India. *Int J Hum Genet*. 2010;10(1–3):1–14.
42. Singh I, Faruq M, Mukherjee O, Jain S, Pal PK, Srivastav MVP, et al. North and South Indian populations share a common ancestral origin of Friedreich's ataxia but vary in age of GAA repeat expansion. *Ann Hum Genet*. 2010; 74(3):202–10.
43. Sachdeva K, Saxena R, Puri R, Bijarnia S, Kohli S, Verma IC. Mutation analysis of the CFTR gene in 225 children: identification of five novel severe and seven reported severe mutations. *Genet Test Mol Biomarkers*. 2012;16(7):798–801.
44. Venugopal A, Chandran M, Eruppakotte N, Kizhakkilach S, Breezevilla SC, Vellingiri B. Monogenic diseases in India. *Mutat Res*. 2018;776:23–31.
45. Verma IC, Bijarnia-Mahay S, Jhingan G, Verma J. Newborn screening: need of the hour in India. *Indian J Pediatr*. 2015;82(1):61–70.
46. Kabra M. Dietary management of inborn errors of metabolism. *Indian J Pediatr*. 2002;69(5):421–6.
47. Sachdeva A. Dietary interventions for rare metabolic disorders—now available in India! *Indian Pediatr*. 2017;54(11):909–10.
48. Mohanty D, Colah RB, Gorakshakar AC, Patel RZ, Master DC, Mahanta J, et al. Prevalence of beta-thalassemia and other haemoglobinopathies in six cities in India: a multicentre study. *J Community Genet*. 2013;4(1):33–42.
49. Nadkarni AH, Gorakshakar AC, Sawant PM, Italia KY, Upadhye DS, Gorivale MS, et al. The phenotypic and molecular diversity of hemoglobinopathies in India: a review of 15 years at a referral center. *Int J Lab Hematol*. 2019;41(2):218–26.
50. Verma IC, Lall M, Dua PR. Down syndrome in India—diagnosis, screening, and prenatal diagnosis. *Clin Lab Med*. 2012;32(2):231–48.
51. Pradhan S, Sengupta M, Dutta A, Bhattacharyya K, Bag SK, Dutta C, et al. Indian genetic disease database. *Nucleic Acids Res*. 2011;39(Database issue):D933–8.
52. Indian Genetic Disease Database. <http://www.igdd.iicb.res.in/>. Accessed 28 Mar 2019.
53. Mahajan A, Chavali S, Kabra M, Chowdhury MR, Bharadwaj D. Molecular characterization of hemophilia B in North Indian families: identification of novel and recurrent molecular events in the factor IX gene. *Haematologica*. 2004;89(12):1498–503.
54. Quadros L, Ghosh K, Shetty S. Novel mutations in factor IX gene from western India with reference to their phenotypic and haplotypic attributes. *J Pediatr Hematol Oncol*. 2009;31(3):157–60.
55. Sharma N, Das R, Kaur J, Ahluwalia J, Trehan A, Bansal D, et al. Evaluation of the genetic basis of phenotypic heterogeneity in north Indian patients with thalassemia major. *Eur J Haematol*. 2010;84(6):531–7.
56. Dash PM, Sahu PK, Patel S, Mashon RS, Kharat KR, Mukherjee MB. Effect of assorted globin haplotypes and alpha-thalassemia on the clinical heterogeneity of Hb S-beta-thalassemia. *Hemoglobin*. 2018;42(4):236–42.
57. Hockham C, Bhatt S, Colah R, Mukherjee MB, Penman BS, Gupta S, et al. The spatial epidemiology of sickle-cell anaemia in India. *Sci Rep*. 2018;8(1):17685.
58. Khan NA, Govindaraj P, Soumitra N, Sharma S, Srilekha S, Ambika S, et al. Leber's hereditary optic neuropathy-specific mutation m.11778G>A exists on diverse mitochondrial haplogroups in India. *Invest Ophthalmol Vis Sci*. 2017;58(10):3923–30.
59. Waldmuller S, Sakthivel S, Saadi AV, Selignow C, Rakesh PG, Golubenko M, et al. Novel deletions in MYH7 and MYBPC3 identified in Indian families with familial hypertrophic cardiomyopathy. *J Mol Cell Cardiol*. 2003;35(6):623–36.
60. Dhandapani PS, Sadayappan S, Xue Y, Powell GT, Rani DS, Nallari P, et al. A common MYBPC3 (cardiac myosin binding protein C) variant associated with cardiomyopathies in South Asia. *Nat Genet*. 2009;41(2):187–91.
61. Lakkakula S, Mohan Pathapati R, Chaubey G, Munirajan AK, Lakkakula BV, Maram R. NAT2 genetic variations among South Indian populations. *Hum genome Var*. 2014;1:14014.
62. Rani DS, Carlus SJ, Poongothai J, Jyothi A, Pavani K, Gupta NJ, et al. CAG repeat variation in the mtDNA polymerase gamma is not associated with oligoasthenozoospermia. *Int J Androl*. 2009;32(6):647–55.
63. Mehrotra S, Oommen J, Mishra A, Sudharshan M, Tiwary P, Jamieson SE, et al. No evidence for association between SLC11A1 and visceral leishmaniasis in India. *BMC Med Genet*. 2011;12:71.
64. Giri AK, Khan NM, Basu A, Tandon N, Scaria V, Bharadwaj D. Pharmacogenetic landscape of clopidogrel in north Indians suggest distinct interpopulation differences in allele frequencies. *Pharmacogenomics*. 2014; 15(5):643–53.
65. Gautam P, Jha P, Kumar D, Tyagi S, Varma B, Dash D, et al. Spectrum of large copy number variations in 26 diverse Indian populations: potential involvement in phenotypic diversity. *Hum Genet*. 2012; 131(1):131–43.
66. Kitzman JO, Mackenzie AP, Adey A, Hiatt JB, Patwardhan RP, Sudmant PH, et al. Haplotype-resolved genome sequencing of a Gujarati Indian individual. *Nat Biotechnol*. 2011;29(1):59–63.
67. Patowary A, Purkanti R, Singh M, Chauhan RK, Bhartiya D, Dwivedi OP, et al. Systematic analysis and functional annotation of variations in the genome of an Indian individual. *Hum Mutat*. 2012;33(7):1133–40.
68. Gupta R, Ratan A, Rajesh C, Chen R, Kim HL, Burhans R, et al. Sequencing and analysis of a South Asian-Indian personal genome. *BMC Genomics*. 2012;13:440.

69. Wong L-P, Lai JK-H, Saw W-Y, Ong RT-H, Cheng AY, Pillai NE, et al. Insights into the genetic structure and diversity of 38 South Asian Indians from deep whole-genome sequencing. *PLoS Genet.* 2014;10(5):e1004377.
70. Chambers JC, Abbott J, Zhang W, Turro E, Scott WR, Tan S-T, et al. The South Asian genome. *PLoS One.* 2014;9(8):e102645.
71. Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO, et al. A global reference for human genetic variation. *Nature.* 2015;526(7571):68–74.
72. Mallick S, Li H, Lipson M, Mathieson I, Gymrek M, Racimo F, et al. The Simons Genome Diversity Project: 300 genomes from 142 diverse populations. *Nature.* 2016;538(7624):201–6.
73. Mondal M, Casals F, Xu T, Dall'Olio GM, Pybus M, Netea MG, et al. Genomic analysis of Andamanese provides insights into ancient human migration into Asia and adaptation. *Nat Genet.* 2016;48(9):1066–70.
74. Rustagi N, Zhou A, Watkins WS, Gedvilaite E, Wang S, Ramesh N, et al. Extremely low-coverage whole genome sequencing in South Asians captures population genomics information. *BMC Genomics.* 2017;18(1):396.
75. Malhotra S, Singh S, Sarkar S. Whole genome variant analysis in three ethnically diverse Indians. *Genes Genomics.* 2018;40(5):497–510.
76. Hariprakash JM, Vellarikkal SK, Verma A, Ranawat BS, Jayarajan R, Ravi R, et al. SAGE: a comprehensive resource of genetic variants integrating South Asian whole genomes and exomes. *Database (Oxford).* 2018;2018:1–10.
77. Almal S, Jeon S, Agarwal M, Patel S, Patel S, Bhak Y, et al. Sequencing and analysis of the whole genome of Indian Gujarati male. *Genomics.* 2019; 111(2):196–204.
78. Ahmed PH, V V, More RP, Viswanath B, Jain S, Rao MS, et al. INDEX-db: The Indian Exome Reference Database (Phase I). *J Comput Biol.* 2019;26(3):225–34.
79. Singh V, Jolly B, Rajput NK, Pramanik S, Bhardwaj A. MtBrowse: An integrative genomics browser for human mitochondrial DNA. *Mitochondrion.* 2019. <https://doi.org/10.1016/j.mito.2019.02.003>
80. Vellarikkal SK, Dhiman H, Joshi K, Hasija Y, Sivasubbu S, Scaria V. mit-o-matic: a comprehensive computational pipeline for clinical evaluation of mitochondrial variations from next-generation sequencing datasets. *Hum Mutat.* 2015;36(4):419–24.
81. Upadhyay P, Gardi N, Desai S, Sahoo B, Singh A, Togar T, et al. TMC-SNPdb: an Indian germline variant database derived from whole exome sequences. *Database (Oxford).* 2016;2016:baw104.
82. An JY. National human genome projects: an update and an agenda. *Epidemiol Health.* 2017;39:e2017045.
83. Richter T, Nestler-Parr S, Babela R, Khan ZM, Tesoro T, Molsen E, et al. Rare disease terminology and definitions—a systematic global review: report of the ISPOR Rare Disease Special Interest Group. *Value Health.* 2015;18(6):906–14.
84. Orphan Drug Act. 1983. <https://www.govinfo.gov/content/pkg/STATUTE-96/pdf/STATUTE-96-Pg2049.pdf>. Accessed 23 May 2019.
85. Thomas S, Caplan A. The Orphan Drug Act Revisited. *JAMA.* 2019;321(9): 833–4.
86. Šimerka P. Council Recommendation. Luxembourg; 2009. <https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:C:2009:151:0007:0010:EN:PDF>. Accessed 28 Mar 2019
87. Passos-Bueno MR, Bertola D, Horovitz DD, de Faria Ferraz VE, Brito LA. Genetics and genomics in Brazil: a promising future. *Mol Genet Genomic Med.* 2014;2(4):280–91.
88. Hsu JC, Wu HC, Feng WC, Chou CH, Lai EC, Lu CY. Disease and economic burden for rare diseases in Taiwan: a longitudinal study using Taiwan's National Health Insurance Research Database. *PLoS One.* 2018;13(9):e0204206.
89. Swaminathan S. ICMR Bulletin. New Delhi; 2017. [https://www.icmr.nic.in/sites/default/files/icmr\\_bulletins/Apr\\_Jun\\_2017.pdf](https://www.icmr.nic.in/sites/default/files/icmr_bulletins/Apr_Jun_2017.pdf). Accessed 28 Ma 2019
90. Rajasimha HK, Shirol PB, Ramamoorthy P, Hegde M, Barde S, Chandru V, et al. Organization for rare diseases India (ORDI)—addressing the challenges and opportunities for the Indian rare diseases' community. *Genet Res (Camb).* 2014;96:e009.
91. Ferreira CR. The burden of rare diseases. *Am J Med Genet A.* 2019;179(6): 885–92.
92. Rare Diseases and Disorders. <http://www.rarediseasesindia.org/>. Accessed 28 Mar 2019.
93. Bhattacharya S, Katoch VM, Majumder PP, Bhattacharya A. Rare diseases In India: current knowledge and new possibilities. *Proc Indian Natl Sci Acad.* 2016;82(4):1183–7.
94. Kar A, Phadnis S, Dharmarajan S, Nakade J. Epidemiology and social costs of haemophilia in India. *Indian J Med Res.* 2014;140(1):19–31.
95. Colah R, Italia K, Gorakshakar A. Burden of thalassemia in India: the road map for control. *Pediatr Hematol Oncol J.* 2017;2(4):79–84.
96. Gourie-Devi M. Epidemiology of neurological disorders in India: review of background, prevalence and incidence of epilepsy, stroke, Parkinson's disease and tremors. *Neurol India.* 2014;62(6):588–98.
97. Molecular Diagnostics, Counseling, Care and Research Centre. <https://www.mdrcindia.org>. Accessed 28 Mar 2019.
98. Bepari KK, Malakar AK, Paul P, Halder B, Chakraborty S. Allele frequency for Cystic fibrosis in Indians vis-a-vis global populations. *Bioinformatics.* 2015; 11(7):348–52.
99. Lakhani R, Ekundayo OT, Shahbazi M. An estimation of the prevalence of intellectual disabilities and its association with age in rural and urban populations in India. *J Neurosci Rural Pract.* 2015;6(4):523–8.
100. Kulkarni ML, Samuel K, Bhagyavathi M, Sureshkumar C. Skeletal dysplasias in a hospital in southern India. *Indian Pediatr.* 1995;32(6):657–65.
101. Singh P, Mukherjee K. Cost-benefit analysis and assessment of quality of care in patients with hemophilia undergoing treatment at National Rural Health Mission in Maharashtra. *India. Value Heal Reg issues.* 2017;12:101–6.
102. Jadhav U, Mukherjee K. Assessment of healthcare measures, healthcare resource use, and cost of care among severe hemophilia A patients in Mumbai region of India. *J Postgrad Med.* 2018;64(3):138–44.
103. Moirangthem A, Phadke SR. Socio-demographic Profile and Economic Burden of Treatment of Transfusion Dependent Thalassemia. *Indian J Pediatr.* 2018;85(2):102–7.
104. Kumar H, Sarma P, Medhi B. Orphan drugs: Indian perspective. *Indian J Pharmacol.* 2017;49:267–9.
105. Indian Rare Disease Registry. <http://bmi.icmr.org.in/irdr/index.php>. Accessed 28 Mar 2019.
106. Phadke S. Hemophilia care in India: a review and experience from a tertiary care centre in Uttar Pradesh. *Indian J Hematol blood.* 2011;27(3):121–6.
107. Jadhav U, Mukherjee K, Lalwani A. Ethical issues in the care of persons living with haemophilia in India. *Indian J Med Ethics.* 2014;11(4):223–7.
108. Hemophilia Federation (India). <http://www.hemophilia.in/index.php/ahf-status>. Accessed 28 Mar 2019.
109. Institute of Medical Genetics and Genomics. <https://www.ncbi.nlm.nih.gov/gtr/labs/217613/>. Accessed 28 Mar 2019.
110. Verma IC, Saxena R, Lall M, Bijarnia S, Sharma R. Genetic counseling and prenatal diagnosis in India—experience at Sir Ganga Ram Hospital. *Indian J Pediatr.* 2003;70(4):293–7.
111. Muranjan M, Karande S. Enzyme replacement therapy in India: lessons and insights. *J Postgrad Med.* 2018;64:195–9.
112. Collins FS, McKusick VA. Implications of the Human Genome Project for medical science. *JAMA.* 2001;285(5):540–4.
113. Wilson BJ, Nicholls SG. The Human Genome Project, and recent advances in personalized genomics. *Risk Manag Healthc Policy.* 2015;8:9–20.
114. Brittain HK, Scott R, Thomas E. The rise of the genome and personalised medicine. *Clin Med.* 2017;17(6):545–51.
115. Chial H. DNA sequencing technologies key to the human genome project. *Nat Educ.* 2008;1(1):219.
116. Heather JM, Chain B. The sequence of sequencers: the history of sequencing DNA. *Genomics.* 2016;107(1):1–8.
117. Shendure J, Balasubramanian S, Church GM, Gilbert W, Rogers J, Schloss JA, et al. DNA sequencing at 40: past, present and future. *Nature.* 2017; 550(7676):345–53.
118. van Dijk EL, Auger H, Jaszczyszyn Y, Thermes C. Ten years of next-generation sequencing technology. *Trends Genet.* 2014;30(9):418–26.
119. Goodwin S, McPherson JD, McCombie WR. Coming of age: ten years of next-generation sequencing technologies. *Nat Rev Genet.* 2016;17(6):333–51.
120. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet.* 2009;10(1):57–63.
121. Necsulea A, Kaessmann H. Evolutionary dynamics of coding and non-coding transcriptomes. *Nat Rev Genet.* 2014;15(11):734–48.
122. Strobel EJ, Yu AM, Lucks JB. High-throughput determination of RNA structures. *Nat Rev Genet.* 2018;19(10):615–34.
123. Meaburn E, Schulz R. Next generation sequencing in epigenetics: insights and challenges. *Semin Cell Dev Biol.* 2012;23(2):192–9.
124. Mias GI, Snyder M. Personal genomes, quantitative dynamic omics and personalized medicine. *Quant Biol.* 2013;1(1):71–90.
125. Chen R, Snyder M. Promise of personalized omics to precision medicine. *Wiley Interdiscip Rev Syst Biol Med.* 2013;5(1):73–82.
126. Need AC, Shashi V, Hitomi Y, Schoch K, Shianna KV, McDonald MT, et al. Clinical application of exome sequencing in undiagnosed genetic conditions. *J Med Genet.* 2012;49(6):353–61.



127. Boycott KM, Vanstone MR, Bulman DE, MacKenzie AE. Rare-disease genetics in the era of next-generation sequencing: discovery to translation. *Nat Rev Genet.* 2013;14(10):681–91.
128. Might M, Wilsey M. The shifting model in clinical diagnostics: how next-generation sequencing and families are altering the way rare diseases are discovered, studied, and treated. *Genet Med.* 2014;16(10):736–7.
129. Lee H, Deignan JL, Dorrani N, Strom SP, Kantarci S, Quintero-Rivera F, et al. Clinical exome sequencing for genetic identification of rare Mendelian disorders. *JAMA.* 2014;312(18):1880–7.
130. Salk JJ, Schmitt MW, Loeb LA. Enhancing the accuracy of next-generation sequencing for detecting rare and subclonal mutations. *Nat Rev Genet.* 2018;19(5):269–85.
131. Rare Disease Impact Report: Insights from patients and the medical community. Shire Human Genetic Therapies. 2013. <https://globalgenes.org/wp-content/uploads/2013/04/ShireReport-1.pdf>. Accessed 23 May 2019.
132. GUAARDIAN. <http://guardian.meragene.com/>. Accessed 28 Mar 2019.
133. Kohler S, Doelken SC, Mungall CJ, Bauer S, Firth HV, Bailleul-Forestier I, et al. The Human Phenotype Ontology project: linking molecular biology and disease through phenotype data. *Nucleic Acids Res.* 2014;42(Database issue):D966–74.
134. Baudhuin LM. Quality guidelines for next-generation sequencing. *Clin Chem.* 2013;59:858–9.
135. Rehm HL, Bale SJ, Bayrak-Toydemir P, Berg JS, Brown KK, Deignan JL, et al. ACMG clinical laboratory standards for next-generation sequencing. *Genet Med.* 2013;15(9):733–47.
136. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. 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.* 2015;17(5):405–24.
137. OpenPGx. <https://web.archive.org/web/20171218064848/http://www.openpgx.org/>. Accessed 23 May 2019.
138. Kandoi G, Nanda A, Scaria V, Sivasubbu S. A case for pharmacogenomics in management of cardiac arrhythmias. *Indian Pacing Electrophysiol J.* 2012;12(2):54–64.
139. Cambon-Thomsen A, Thorisson GA, Mabille L. The role of a bioresource research impact factor as an incentive to share human bioresources. *Nat Gen.* 2011;43:503–4.
140. Vellarikkal SK, Patowary A, Singh M, Kumari R, Faruq M, Master DC, et al. Exome sequencing reveals a novel mutation, p.L325H, in the KRT5 gene associated with autosomal dominant Epidermolysis Bullosa Simplex Koebner type in a large family from western India. *Hum genome Var.* 2014;1:14007.
141. Karuthedath Vellarikkal S, Jayarajan R, Verma A, Nair S, Ravi R, Senthilvel V, et al. Case Report: Whole exome sequencing reveals a novel frameshift deletion mutation p.G2254fs in COL7A1 associated with autosomal recessive dystrophic epidermolysis bullosa. *F1000Res.* 2016;5:900.
142. Yenamandra VK, Vellarikkal SK, Kumar M, Chowdhury MR, Jayarajan R, Verma A, et al. Application of whole exome sequencing in elucidating the phenotype and genotype spectrum of junctional epidermolysis bullosa: a preliminary experience of a tertiary care centre in India. *J Dermatol Sci.* 2017;86(1):30–6.
143. Yenamandra VK, Vellarikkal SK, Chowdhury MR, Jayarajan R, Verma A, Scaria V, et al. Genotype-phenotype correlations of dystrophic epidermolysis bullosa in India: experience from a tertiary care centre. *Acta Derm Venereol.* 2018;98(9):873–9.
144. Sandhya P, Vellarikkal SK, Nair A, Ravi R, Mathew J, Jayarajan R, et al. Egyptian tale from India: application of whole-exome sequencing in diagnosis of atypical familial Mediterranean fever. *Int J Rheum Dis.* 2017;20(11):1770–5.
145. Gupta A, Sharma Y, Deo K, Vellarikkal S, Jayarajan R, Dixit V, et al. Case Report: Whole exome sequencing helps in accurate molecular diagnosis in siblings with a rare co-occurrence of paternally inherited 22q12 duplication and autosomal recessive non-syndromic ichthyosis. *F1000Res.* 2015;4:446.
146. Gupta A, Sharma YK, Vellarikkal SK, Jayarajan R, Dixit V, Verma A, et al. Whole-exome sequencing solves diagnostic dilemma in a rare case of sporadic acrokeratosis verruciformis. *J Eur Acad Dermatol Venereol.* 2016;30:695–7.
147. Narayanan R, Karuthedath Vellarikkal S, Jayarajan R, Verma A, Dixit V, Scaria V, et al. Case report: application of whole exome sequencing for accurate diagnosis of rare syndromes of mineralocorticoid excess. *F1000Res.* 2016;5:1592.
148. Govindaraj GM, Karuthedath Vellarikkal S, Jayarajan R, Ravi R, Verma A, Chakkiyar K, et al. Case Report: Whole exome sequencing identifies variation c.2308G>A p.E770K in RAG1 associated with B- T- NK+ severe combined immunodeficiency. *F1000Res.* 2016;5:2532.
149. Rawat A, Karuthedath Vellarikkal S, Verma A, Jayarajan R, Gupta A, Singh S, et al. Case Report: Whole exome sequencing identifies a novel frameshift insertion c.1325dupT (p.F442fsX2) in the tyrosine kinase domain of BTK gene in a young Indian individual with X-linked agammaglobulinemia. *F1000Res.* 2016;5:2667.
150. Mg G, Riyaz A, Krishnan C, Scaria V. Rapid transition of facial features from early to mid-adolescence in autosomal dominant hyper IgE syndrome with a STAT3 Variation. *Indian J Pediatr.* 2018;85:595–6.
151. Virmani N, Vellarikkal SK, Verma A, Jayarajan R, Sakhiya J, Desai C, et al. Whole exome sequencing in a multi-generation family from India reveals a genetic variation c.10C>T (p.Gln4Ter) in keratin 5 gene associated with Dowling-Degos disease. *Indian J Dermatol Venereol Leprol.* 2018;84:344–6.
152. Vellarikkal SK, Jayarajan R, Verma A, Ravi R, Senthilvel V, Kumar A, et al. A founder mutation MLC1 c.736delA associated with megalencephalic leukoencephalopathy with subcortical cysts-1 in north Indian kindred. *Clin Genet.* 2018;94:271–3.
153. Yenamandra VK, Shamsudheen KV, Madhumita RC, Rijith J, Ankit V, Scaria V, et al. Autosomal recessive epidermolysis bullosa simplex: report of three cases from India. *Clin Exp Dermatol.* 2017;42:800–3.
154. Mahajan R, Vellarikkal SK, Handa S, Verma A, Jayarajan R, Kumar A, et al. Utility of whole-exome sequencing in detecting novel compound heterozygous mutations in COL7A1 among families with severe recessive dystrophic epidermolysis bullosa in India - implications on diagnosis, prognosis and prenatal testing. *J Eur Acad Dermatol Venereol.* 2018;32:e433–5.
155. Hariprakash JM, Vellarikkal SK, Keechilat P, Verma A, Jayarajan R, Dixit V, et al. Pharmacogenetic landscape of DPYD variants in south Asian populations by integration of genome-scale data. *Pharmacogenomics.* 2018;19(3):227–41.
156. Sivasdas A, Salleh MZ, Teh LK, Scaria V. Genetic epidemiology of pharmacogenetic variants in South East Asian Malays using whole-genome sequences. *Pharmacogenomics J.* 2017;17(5):461–70.
157. Sivasdas A, Sharma P, Scaria V. Landscape of warfarin and clopidogrel pharmacogenetic variants in Qatari population from whole exome datasets. *Pharmacogenomics.* 2016;17(17):1891–901.
158. Sivasdas A, Scaria V. Pharmacogenomic survey of Qatari populations using whole-genome and exome sequences. *Pharmacogenomics J.* 2018;18(4):590–600.
159. Koshy R, Ranawat A, Scaria V. al mena: a comprehensive resource of human genetic variants integrating genomes and exomes from Arab, Middle Eastern and North African populations. *J Hum Genet.* 2017;62(10):889–94.
160. Jain A, Gandhi S, Koshy R, Scaria V. Incidental and clinically actionable genetic variants in 1005 whole exomes and genomes from Qatar. *Mol Genet Genomics.* 2018;293(4):919–29.
161. Scaria V, Sivasubbu S. Exome Sequence Analysis and Interpretation: Handbook for Clinicians. 1st ed: The Readers Paradise; 2015. p. 200.
162. Enns GM, Shashi V, Bainbridge M, Gambello MJ, Zahir FR, Bast T, et al. Mutations in NGLY1 cause an inherited disorder of the endoplasmic reticulum-associated degradation pathway. *Genet Med.* 2014;16(10):751–8.
163. Abouelhoda M, Sobahy T, El-Kalioby M, Patel N, Shamseldin H, Monies D, et al. Clinical genomics can facilitate countrywide estimation of autosomal recessive disease burden. *Genet Med.* 2016;18(12):1244–9.
164. Kumari P, Joshi D, Shamal SN, Singh R. Study of dystrophinopathy in Eastern Uttar Pradesh population of India. *J Pediatr Neurosci.* 2018;13(2):182–8.
165. Basumatary LJ, Das M, Goswami M, Kayal AK. Deletion pattern in the dystrophin gene in Duchenne muscular dystrophy patients in northeast India. *J Neurosci Rural Pract.* 2013;4(2):227–9.
166. Singh R-J, Manjunath M, Preethish-Kumar V, Polavarapu K, Vengalil S, Thomas PT, et al. Natural history of a cohort of Duchenne muscular dystrophy children seen between 1998 and 2014: An observational study from South India. *Neurol India.* 2018;66(1):77–82.
167. Sakthivel Murugan SM, Arthi C, Thilothammal N, Lakshmi BR. Carrier detection in Duchenne muscular dystrophy using molecular methods. *Indian J Med Res.* 2013;137(6):1102–10.
168. Vengalil S, Preethish-Kumar V, Polavarapu K, Mahadevappa M, Sekar D, Purushottam M, et al. Duchenne muscular dystrophy and Becker muscular dystrophy confirmed by multiplex ligation-dependent probe amplification: genotype-phenotype correlation in a large cohort. *J Clin Neurol.* 2017;13(1):91–7.

169. Manjunath M, Kiran P, Preethish-Kumar V, Nalini A, Singh RJ, Gayathri N. A comparative study of mPCR, MLPA, and muscle biopsy results in a cohort of children with Duchenne muscular dystrophy: a first study. *Neurol India*. 2015;63(1):58–62.
170. Deeptha S, Vengalil S, Preethish-Kumar V, Polavarapu K, Nalini A, Gayathri N, et al. MLPA identification of dystrophin mutations and in silico evaluation of the predicted protein in dystrophinopathy cases from India. *BMC Med Genet*. 2017;18(1):67.
171. Singh B, Mandal K, Lallar M, Narayanan DL, Mishra S, Gambhir PS, et al. Next generation sequencing in diagnosis of MLPA negative cases presenting as Duchenne/Becker muscular dystrophies. *Indian J Pediatr*. 2018;85:309–10.
172. Kadali S, Kolusu A, Gummadi MR, Undamatla J. The relative frequency of lysosomal storage disorders: a medical genetics referral laboratory's experience from India. *J Child Neurol*. 2014;29(10):1377–82.
173. Hebbar M, Prasada LH, Bhowmik A, Trujillano D, Shukla A, Chakraborti S, et al. Homozygous deletion of exons 2 and 3 of NPC2 associated with Niemann-Pick disease type C. *Am J Med Gen Part A*. 2016;170:2486–9.
174. Madkaikar M, Aluri J, Gupta S. Guidelines for screening, early diagnosis and management of severe combined immunodeficiency (SCID) in India. *Indian J Pediatr*. 2016;83(5):455–62.
175. Jindal AK, Pilania RK, Rawat A, Singh S. Primary immunodeficiency disorders in India—a situational review. *Front Immunol*. 2017;8:714.
176. Gupta M, Aluri J, Desai M, Lokeshwar M, Taur P, Lenardo M, et al. Clinical, immunological, and molecular findings in four cases of B cell expansion with NF-kappaB and T cell anergy disease for the first time from India. *Front Immunol*. 2018;9:1049.
177. Aluri J, Desai M, Gupta M, Dalvi A, Terance A, Rosenzweig SD, et al. Clinical, immunological, and molecular findings in 57 patients with severe combined immunodeficiency (SCID) from India. *Front Immunol*. 2019;10:23.
178. Aluri J, Gupta M, Dalvi A, Mhatre S, Kulkarni M, Hule G, et al. Clinical, immunological, and molecular findings in five patients with major histocompatibility complex class II deficiency from India. *Front Immunol*. 2018;9:188.
179. Sonam K, Bindu PS, Srinivas Bharath MM, Govindaraj P, Gayathri N, Arvinda HR, et al. Mitochondrial oxidative phosphorylation disorders in children: phenotypic, genotypic and biochemical correlations in 85 patients from South India. *Mitochondrion*. 2017;32:42–9.
180. Vinu N, Puri RD, Anand K, Verma IC. Expanding the phenotype of the founder South Asian mutation in the nuclear encoding mitochondrial RMND1 Gene. *Indian J Pediatr*. 2018;85(2):87–92.
181. Srivastava A, Srivastava KR, Hebbar M, Galada C, Kadavigrere R, Su F, et al. Genetic diversity of NDUFV1-dependent mitochondrial complex I deficiency. *Eur J Hum Genet*. 2018;26(11):1582–7.
182. Dalal A, Das BA, Agarwal D, Phadke SR. Exome sequencing & homozygosity mapping for identification of genetic aetiology for spastic ataxia in a consanguineous family. *Indian J Med Res*. 2015;142:220–4.
183. Faruq M, Narang A, Kumari R, Pandey R, Garg A, Behari M, et al. Novel mutations in typical and atypical genetic loci through exome sequencing in autosomal recessive cerebellar ataxia families. *Clin Genet*. 2014;86(4):335–41.
184. Uttarilli A, Shah H, Bhavani GS, Upadhyai P, Shukla A, Girisha KM. Phenotyping and genotyping of skeletal dysplasias: evolution of a center and a decade of experience in India. *Bone*. 2019;120:204–11.
185. Girisha KM, Kortum F, Shah H, Alawi M, Dalal A, Bhavani GS, et al. A novel multiple joint dislocation syndrome associated with a homozygous nonsense variant in the EXOC6B gene. *Eur J Hum Genet*. 2016;24(8):1206–10.
186. Das Bhowmik A, Dalal A, Matta D, Kandadai RM, Kanikannan MA, Aggarwal S. Identification of a novel splice site HSPG2 mutation and prenatal diagnosis in Schwartz Jampel Syndrome type 1 using whole exome sequencing. *Neuromuscul Disord*. 2016;26(11):809–14.
187. Rai E, Mahajan A, Kumar P, Angural A, Dhar MK, Razdan S, et al. Whole exome screening identifies novel and recurrent WISP3 mutations causing progressive pseudorheumatoid dysplasia in Jammu and Kashmir-India. *Sci Rep*. 2016;6:27684.
188. Genomics and other Omics tools for Enabling Medical Decision. <http://gomedigib.in/>. Accessed 28 Mar 2019.
189. Kar B, Sivamani S. Directory of genetic test services and counselling centres in India. *Int J Hum Genet*. 2016;16(3–4):148–57.
190. India Newborn Action Plan - National Health Mission. <http://nhm.gov.in/images/pdf/programmes/inap-final.pdf>. Accessed 28 Mar 2019.
191. Chakrabarty S, Kabekkodu SP, Brand A, Satyamoorthy K. Perspectives on translational genomics and public health in India. *Public Health Genomics*. 2016;19(2):61–8.
192. Puri RD, Kabra M. Editorial: new horizons in genetic diagnosis in pediatric practice: the excitement and challenges! *Indian J Pediatr*. 2016;83:1131–2.
193. National Policy For Treatment Of Rare Diseases. 2017. [https://mohfw.gov.in/sites/default/files/Rare Diseases Policy FINAL.pdf](https://mohfw.gov.in/sites/default/files/Rare%20Diseases%20Policy%20FINAL.pdf). Accessed 28 Mar 2019.
194. Rare disease cell order. 2018. <https://mohfw.gov.in/sites/default/files/National-policy-for-Treatment-of-Rare-Diseases.pdf>. Accessed 28 Mar 2019.
195. Rivas MA, Avila BE, Koskela J, Huang H, Stevens C, Pirinen M, et al. Insights into the genetic epidemiology of Crohn's and rare diseases in the Ashkenazi Jewish population. *PLoS Genet*. 2018;14(5):e1007329.
196. Genome Asia 100K. <http://www.genomeasia100k.com/>. Accessed 28 Mar 2019.
197. First PM-STIAC Meet. [http://psa.gov.in/archive/pmstiac\\_first\\_meeting](http://psa.gov.in/archive/pmstiac_first_meeting). Accessed 28 Mar 2019.
198. IndiGen 2019. <https://indigen.igib.in/>. Accessed 23 May 2019.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

