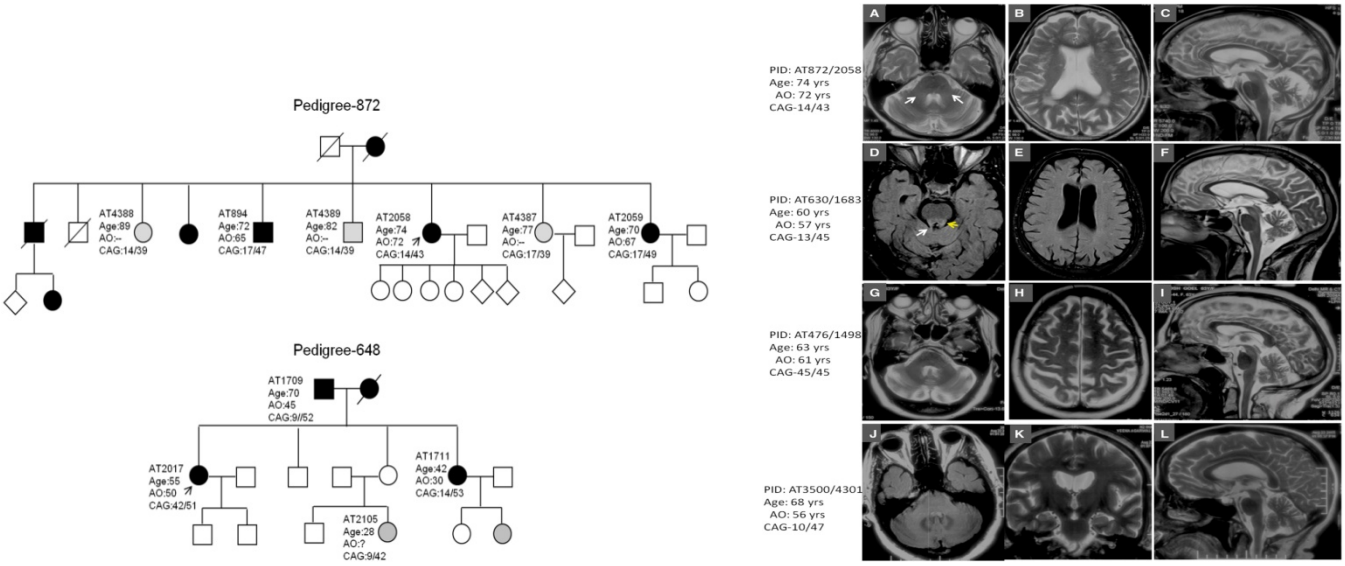
10 Best publication and details

A. Work done in the area of Hereditary ataxia and related neurogenetic disorders

1. ***Clinical behaviour of spinocerebellar ataxia type 12 and intermediate length abnormal CAG repeats in PPP2R2B.***
   Brain. 2017 *Jan;140(1):27-36.*
   Srivastava AK, Takkar A, Garg A, ***Faruq M****.*

Highlights: This manuscript described the definition of 'The pathogenic length' of CAG repeat expansion in PPP2R2B gene that is the causal factor for SCA12. This was established by taking into account of the clinical phenotypes of wide varieties of CAG lengths determined in SCA12 patients, a unique subtype of SCA in Indian population. A detailed analysis described the clinical behaviour of 18 patients who harbour CAG repeats in the range of 43-50 and compare their clinical behaviour with patients carrying typical pathogenic threshold length of 51 CAG repeats. It was observed that the clinical characteristics were similar to those of typical SCA12 phenotype, with large variability in the age at onset. Radiologically, we observed a variable degree of cerebro-cerebellar degeneration along with white matter changes that do not correlate with the disease severity. ***We define a new pathogenic threshold of CAG-43 to be pathogenic for SCA12 diagnosis*** and also describe the clinical profiles of two biallelic CAG expansion carriers. We also proposed that SCA12 might not be that restricted in terms of occurrence in other geographical or ethnic populations, as it was previously presumed to be.



Pedigree chart two families with rarest presentation of i) Smallest Pathogenic length of SCA12 CAG repeats ii) Biallelic CAG expansion, and MRI images depcting various brain parenchymal changes

2. *Generation of three spinocerebellar ataxia type-12 patients derived induced pluripotent stem cell lines (IGIBi002-A,IGIBi003-A and IGIBi004-A).*
Stem Cell Res. 2018 *Aug;31:216-221*
Kumar D, Hussain A, Srivastava AK, Mukerji M, Mukherjee O, ***Faruq M****.*

Highlights: This is the ***first report of generation of Induced Pluripotent stem cell lines from blood cells of SCA12*** patients (n=3). These iPSc lines lines were used for differentiation into neurons.

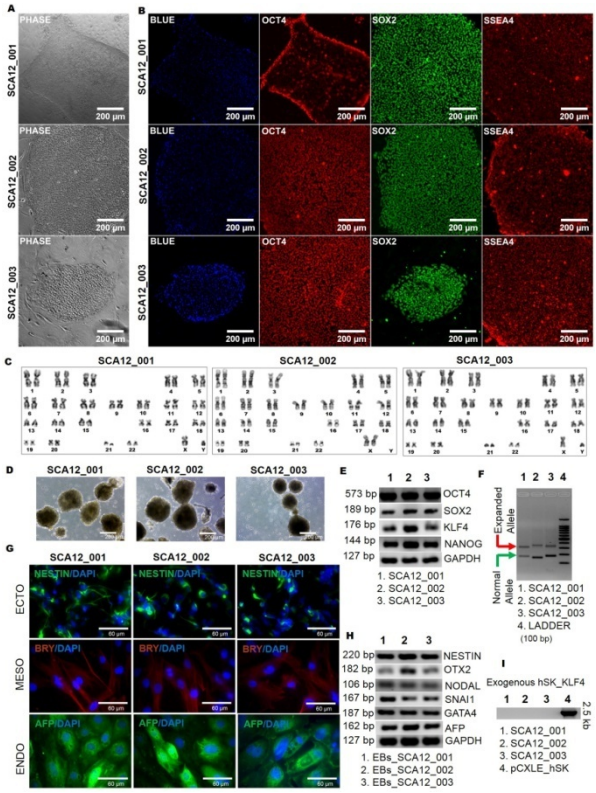


Fig. 1. Generation and characterization of iPSC lines (IGIBi002-A, IGIBi003-A and IGIBi004-A)

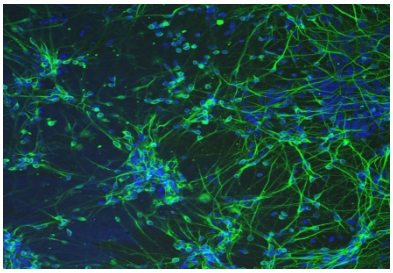


Fig..2. Differentiated Neuronal lineage from a patient line AT-2555 Neuron labelled with MAP2

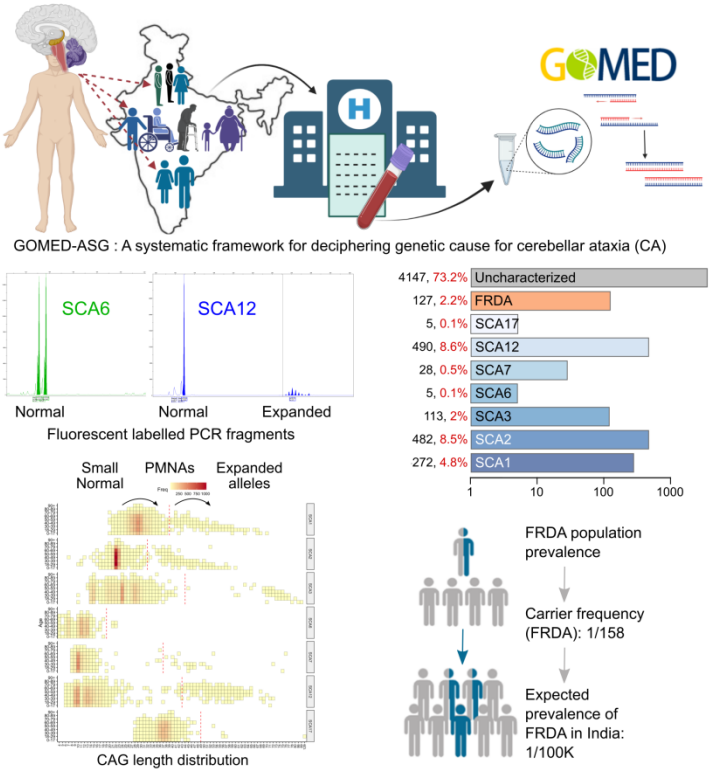
3. Genetics of Ataxias in Indian Population: A Collative Insight from a Common Genetic Screening Tool,

Advanced Genetics, *Advanced genetics (Hoboken, N.J.)*, *3*(2), 2100078

Pooja Sharma, Akhilesh Kumar Sonakar, Nishu Tyagi, Varun Suroliya, Manish Kumar, Rintu Kutum, Vivekananda Asokchandran, Sakshi Ambawat, Uzma Shamim, Avni Anand, Ishtaq Ahmad, Sunil Shakya, Bharathram Uppili, Aradhana Mathur, Shaista Parveen, Shweta Jain, Jyotsna Singh, Malika Seth, Sana Zahra, Aditi Joshi, Divya Goel, Shweta Sahni, Asangla Kamai, Saruchi Wadhwa, Aparna Murali, Sheeba Saifi, Debashish Chowdhury, Sanjay Pandey, Kuljeet Singh Anand, Ranganathan Lakshmi Narasimhan, Sanghamitra Laskar, Suman Kushwaha, Mukesh Kumar, Cheruvallill Velayudhan Shaji, Madakasira Vasantha Padma Srivastava, Achal K. Srivastava, **Mohammed Faruq***,

Highlights: Here, with this study, we shared our experience from investigations of SCAs linked to Tandem nucleotide repeat expansion mutations in the ***largest cohort (~5600) till date from India.*** This study defines that 30-40% of ataxia cohort (Indian) is explained by just 6 different SCA subtypes (Autosomal dominant types: SCA1, SCA2, SCA3, SCA6, SCA7, and SCA12) and Friedreich ataxia (a recessive form of ataxia). Further, we have analyzed the occurrence of SCAs, prevalent mutations and attempted to assess the following points:

- Overall differential distribution of SCAs in India across the Institutes.
- Diagnostic yield of each SCA subtype across different centre.
- Sensitivity of clinical diagnosis.
- Occurrence of pre-mutation alleles for each of the SCA mutation type and other relevant observations



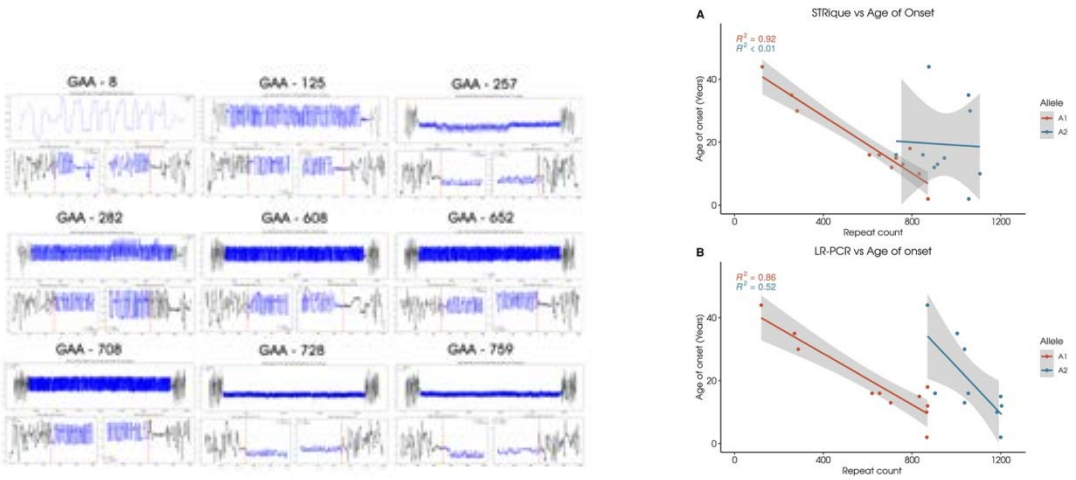
GOMED-ASG : A systematic framework for deciphering genetic cause for cerebellar ataxia (CA)

4. Sequencing through hyperexpanded Friedreich's ataxia-GAA repeats by nanopore technology: implications in genotype–phenotype correlation,

Brain Communications, Volume 5, Issue 2, 2023,

Bharathram Uppili, Pooja Sharma, Istaq Ahmad, Shweta Sahni, Vivekanand Asokachandran, Anil B Nagaraja, Achal K Srivastava, **Mohammed Faruq**,

Friedreich's ataxia, an autosomal recessive disorder, is caused by tandem GAA nucleotide repeat expansions in intron 1 of the frataxin gene. The GAA repeats over 66 in number are considered as pathogenic, and commonly occurring pathogenic repeats are within a range of 600–1200. Clinically, the spectrum of features is confined mainly to neurological tissues; however, cardiomyopathy and diabetes mellitus have been reported in 60 and 30% of the subjects, respectively. ***The accurate detection of GAA repeat count is of utmost importance for clinical genetic correlation, and no study so far has attempted an approach that is of high-throughput nature and defines the exact sequence of GAA repeats.*** Largely, the method for detection of GAA repeats so far is either through the conventional polymerase chain reaction-based screening or Southern blot, which remains the gold standard method. We utilized an approach of long-range targeted amplification of *FXN*-GAA repeats using Oxford Nanopore Technologies MinION platform for accurate estimation of repeat length. We were able to achieve successful amplification of GAA repeats ranging from ~120 to 1100 at ~2600× mean coverage. The total throughput achievable through our protocol can allow for screening of up to 96 samples per flow cell in less than 24 h. The proposed method is clinically scalable and deployable for day-to-day diagnostics. In this paper, we demonstrate to resolve the genotype–phenotype correlation of Friedreich's ataxia patients with better accuracy.



Left; Squiggles representative of GAA repeat stretch on Nanopore sequencing data
Right: Improved Age at onset versus GAA repeat correlation coefficient with GAA sequencing. Correlation between age of onset and allelic length as estimated by different methods (**A**) STRique and (**B**) RP-PCR.

5. **Whole exome and targeted gene sequencing todetect pathogenic recessive variants in early onset cerebellar ataxia.**
Clin Genet. 2019 Dec;96(6):566-574.
Shakya S, Kumari R, Suroliya V, Tyagi N, Joshi A, Garg A, Singh I, Kalikavil Puthanveedu D, Cherian A, Mukerji M, Srivastava AK, **Faruq M**.

Highlights: This is the second exploratory study utilizing NGS based tools to decipher genetics of ataxias beyond tandem nucleotide repeats. As explained in previous study 30-40% of heterditary ataxias are caused by Tandem nucleotide repeats, the remaining may alternate mutation mechanism (point mutations etc) and there are nearly 100 different genetic etiologies that can explain the cause. Therefore to decipher the remaining cases, we undertook this study that utilized the power of NGS and whole exome sequencing followed by targeted panel sequencing lead to the identification of rare yet first description of those ataxia types in Indian population.

A total of 98 index patients were clinically and genetically (whole exome sequencing (WES) in 16 patients and targeted gene panel of 41 ataxia causing genes in 82 patients) evaluated. Four families underwent WES in family based design. Overall, we have identified 24 variants comprising 20 pathogenic and four likely-pathogenic both rare/novel, variations in 21 early onset cerebellar ataxia patients. Among the identified variations, SACS (Spastic ataxia of Charlex saguinay type, n = 7) and SETX (Oculomotor Apraxia type-2, n = 6) were frequent, while ATM (Ataxia Telengiectasia, n = 2), TTPA (Vitamin E gene Deficiency; n = 2) and other rare loci were observed. We have prioritized novel pathogenic variants in RARS2 and FA2H loci through family based design in two out of four families.

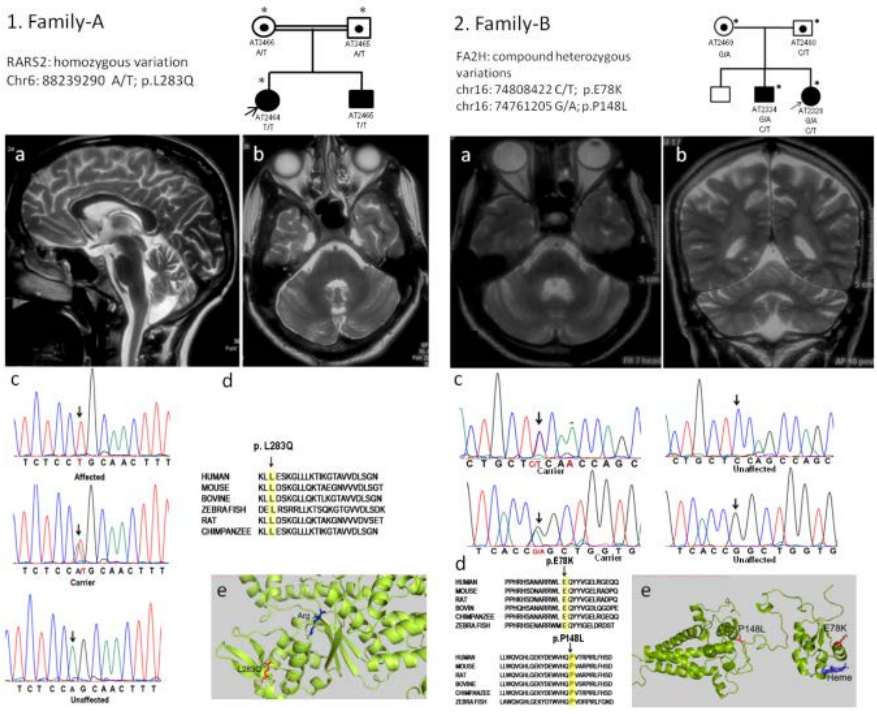


Figure showing Two families with rarest form of ataxia types caused by RARS2 and FA2H mutation.

6. Synthetic transcription elongation factors license transcription across repressive chromatin

Science. **2017** Dec 22;358(6370):1617-1622.

Erwin GS, Grieshop MP, Ali A, Qi J, Lawlor M, Kumar D, Ahmad I, McNally A, Teider N, Worringer K, Sivasankaran R, Syed DN, Eguchi A, Ashraf M, Jeffery J, Xu M, Park PMC, Mukhtar H, Srivastava AK, Faruq M, Bradner JE, Ansari AZ.

This study is a major landmark research where a synthetic molecule Syn-TEF1 (Transcription Elongation factor), a molecule that actively enables transcription across repressive GAA repeats that silence frataxin expression in Friedreich's ataxia, a terminal neurodegenerative disease with no effective therapy. The modular design of Syn-TEF1 defines a general framework for developing a class of molecules that license transcription elongation at targeted genomic loci.

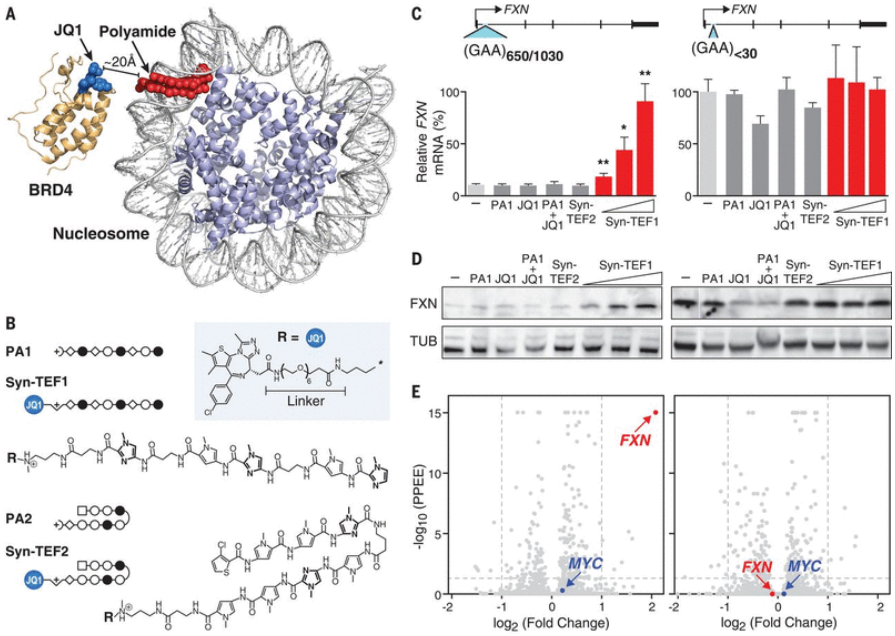


Figure: Synthetic transcription elongation factors (Syn-TEFs) selectively activate FXN expression (A) Co-crystal structures of JQ1 bound to BRD4 [Protein Data Bank (PDB) ID, 3MXF] and polyamide bound to nucleosomal DNA (PDB ID, 1M1A). (B) Linear PA1 and Syn-TEF1 target the DNA sequence 5′- AAGAAGAAG-3′. Hairpin PA2 and Syn-TEF2 target 5′-WTACGTW-3′, where W is A or T. (C) Relative expression of FXN mRNA in GM15850 (left) and GM15851 (right) cell lines by quantitative reverse transcription polymerase chain reaction. Results are means ± SEM (n = 4), normalized to the relative expression of FXN in GM15851 cells (fig. S4). (D) Immunoblot of FXN and a-tubulin (TUB) with treated GM15850 (left) and GM15851 (right) cells. Cells were treated as in (C). (E) Volcano plots of RNA-seq data display the change in global gene expression after 24 hours of treatment of GM15850 (left) and GM15851 (right) cells with 1 mM Syn-TEF1 (n = 4).

Here in this study, our patient derived disease cell line models of LCLs and primary cell lines resource from FRDA patients helped in validating and substantiating the effect of Syn-TEFs.

7. **C9orf72 hexanucleotide repeat expansion in Indian patients with ALS: a common founder and its geographical predilection.**

Neurobiol Aging. 2020 Jan 3. pii: S0197-4580(19)30453-1.

Shamim U, Ambawat S, Singh J, Thomas A, Pradeep-Chandra-Reddy C, Suroliya V, Uppilli B, Parveen S, Sharma P, Chanchal S, Nashi S, Preethish-Kumar V, Vengalil S, Polavarapu K, Keerthipriya M, Mahajan NP, Reddy N, Thomas PT, Sadasivan A, Warrier M, Seth M, Zahra S,Mathur A, Vibha D, Srivastava AK, Nalini A, **Faruq M***.

Highlights: In this study we presented the largest series of Amyotrophic Lateral Sclerosis patients investigated for Hexanucleotide repeats G4C2(GGGGCC)n. We assessed its frequency in a study cohort involving 593 clinically and electrophysiologically defined ALS cases. We also investigated the presence of reported shared haplotype among the mutation carriers. The G4C2 expansion was observed in 3.2% (19/593) of total cases where 9/19 (47.4%) positive cases belonged to the eastern region of India.

Haplotype analysis revealed 11 G4C2-Ex carriers shared the common haplotype (haplo-A) background spanning a region of ~90 kbp (rs895021-rs11789520) including rs3849942 (a well-known global at-risk loci with T allele for G4C2 expansion). The other 3 G4C2-Ex cases had a different haplotype (haplo-B) with core difference from haplo-A at G4C2-Ex flanking 31 kbp region between rs3849942 and rs11789520 SNPs (allele 'C' of rs3849942 which is a non-risk allele).

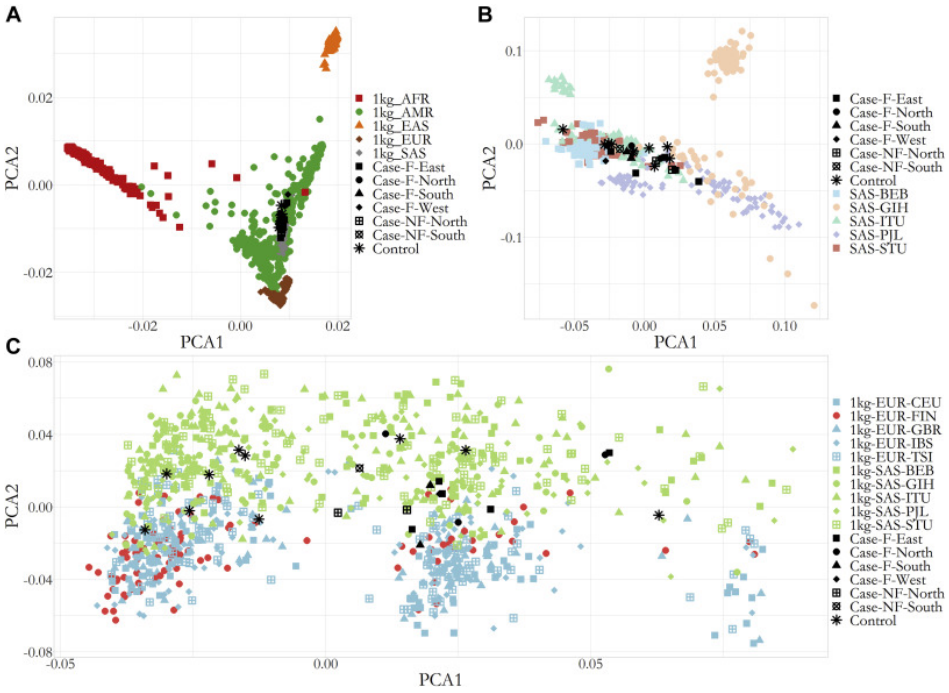


Figure: PCA showing the genetic similarities of c9ORF72 positive cases with European population.

B. Major work done in the area of Clinical Genetic and Genomics

8. Biallelic cGMP-dependent type II protein kinase gene (*PRKG2*) variants cause a novel acromesomelic dysplasia.

J. Med Genet. 2020 Oct 26:jmedgenet-2020-107177. doi: 10.1136/jmedgenet-2020-107177.

Díaz-González F, Wadhwa S, Rodriguez-Zabala M, Kumar S, Aza-Carmona M, Sentchordi-Montané L, Alonso M, Ahmad I, Zahra S, Kumar D, Kushwah N, Shamim U, Sait H, Kapoor S, Roldán B, Nishimura G, Offiah AC, **Faruq M**, Heath KE.J.

Highlights: ***This study report for the first time about a novel Skeletal Dysplasia type ACromesomelic Dysplasia caused by PRKG2 (cGMP-dependent type II protein kinase) gene***. We described in equal collaboration two patients with this unique clinical phenotype, one Indian patient and another of Spanish origin. The clinical features broadly categorized as severe short stature due to acromesomelic limb shortening, brachydactyly, mild to moderate platyspondyly and progressively increasing metaphyseal alterations of the long bones. Functional characterisation was undertaken for the identified frameshift variants. We further demonstrated that the mutant transcripts are exposed to nonsense-mediated decay and the truncated mutant cGKII proteins, partially or completely lacking the kinase domain, alter the downstream mitogen activation protein kinase signalling pathway by failing to phosphorylate c-Raf 1 at Ser43 and subsequently reduce ERK1/2 activation in response to fibroblast growth factor 2. They also downregulate *COL10A1* and upregulate *COL2A1* expression through SOX9.

Figure- Xray depicting major skeletal abnormalities.

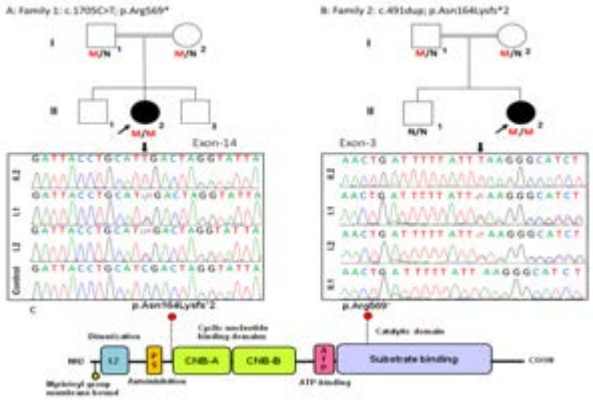


Illustration showing mutant PRKG2

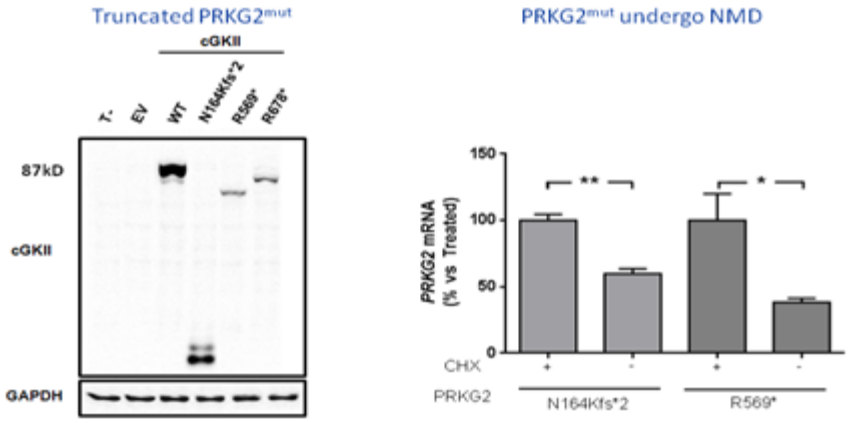


Figure: showing the effect of Framshift variations leading to premature truncation and non-sense mediated decay due to framshift nature of the mutation.

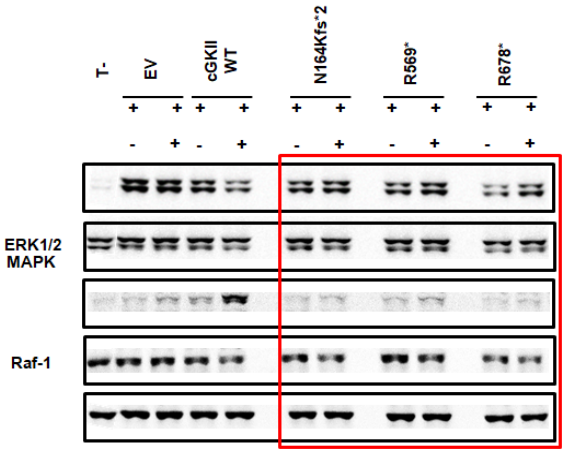


Figure: PRKG2 mutants are unable to down-regulate MAPK signaling
Neither human PRKG2 mutants, nor the positive control bovine mutant are able to downregulate ERK activation

An iPSc lne of this patient of Indian origin was generated in my lab.
****Kumar M, Wadhwa S, Tyagi N, Ahmad I, Kumar S, Sagar S, Zahra S, Kamai A, Shamim U, Kapoor S, Faruq M. **Generation of induced pluripotent stem cell line (IGIBi007-A) from a patient with a novel acromesomelic dysplasia, PRKG2 type (AMDP).** Stem Cell Res. 2021 May;53:102340. doi: 10.1016/j.scr.2021.102340. Epub 2021 Apr 19. PMID: 33887582.

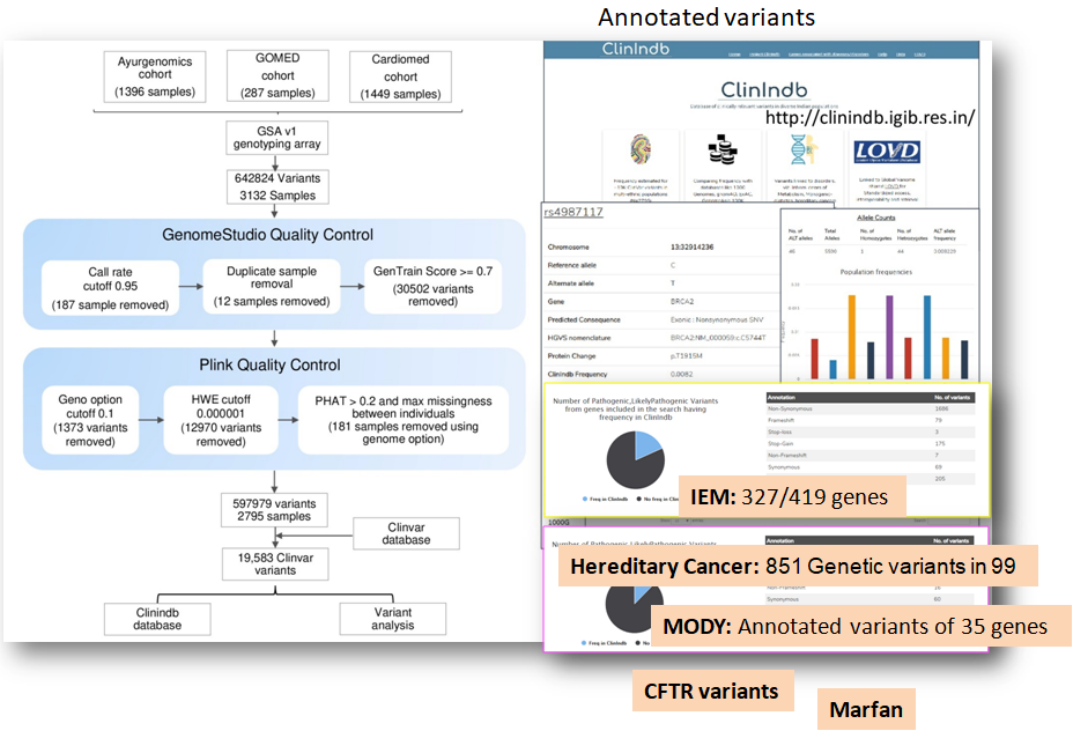
9. Frequency spectrum of rare and clinically relevant markers in multiethnic Indian populations (ClinIndb): A resource for genomic medicine in India.

Hum Mutat. 2020 Nov;41(11):1833-1847.

Narang A, Uppilli B, Vivekanand A, Naushin S, Yadav A, Singhal K, Shamim U, Sharma P, Zahra S, Mathur A, Seth M, Parveen S, Vats A, Hillman S, Dolma P, Varma B, Jain V; TRISUTRA Ayurgenomics Consortium, Prasher B, Sengupta S, Mukerji M, **Faruq M*.**

The aim of this study was to estimate the burden of monogenic disease-causing variants in Indian populations. Toward this, we have assessed the frequency profile of monogenic phenotype-associated ClinVar variants. The study utilized a genotype data set (global screening array, Illumina) from 2795 individuals (multiple in-house genomics cohorts) representing diverse ethnic and geographically distinct Indian populations. Of the analyzed variants from Global Screening Array, ~9% were found to be informative and were either not known earlier or underrepresented in public databases in terms of their frequencies. These variants were linked to disorders, namely inborn errors of metabolism, monogenic diabetes, hereditary cancers, and various other hereditary conditions. We have also shown that our study cohort is genetically a better representative of the Indian population than its representation in the 1000 Genome Project (South Asians). We have created a database, ClinIndb, linked to the Leiden Open Variation Database, to help clinicians and researchers in diagnosis, counseling, and development of appropriate genetic screening tools relevant to the Indian populations and Indians living abroad.

ClinIndb- Disease Allele Frequency resource for Indian population



C. Work done about COVID genome surveillance

10. Integrated genomic view of SARS-CoV-2 in India.

Wellcome Open Res. 2020 Aug 3;5:184. doi: 10.12688/wellcomeopenres.16119.1. eCollection 2020.PMID: 32995557

Kumar P, Pandey R, Sharma P, Dhar MS, A V, Uppili B, Vashisht H, Wadhwa S, Tyagi N, Fatihi S, Sharma U, Singh P, Lall H, Datta M, Gupta P, Saini N, Tewari A, Nandi B, Kumar D, Bag S, Gahlot D, Rathore S, Jatana N, Jaiswal V, Gogia H, Madan P, Singh S, Singh P, Dash D, Bala M, Kabra S, Singh S, Mukerji M, Thukral L, **Faruq M***, Agrawal A, Rakshit P.

Highlights: This work was lead by my lab in collaboration with NCDC (National Centre for Disease Control). ***This work is the earliest description of 104 COVID genomes sequenced from Delhi and surrounding region during April-May 2020***. The analyses revealed multiple introductions of SARS-CoV-2 genomes, including the A2a cluster from Europe and the USA, A3 cluster from Middle East and A4 cluster (haplotype redefined) from Southeast Asia (Indonesia, Thailand and Malaysia) and Central Asia (Kyrgyzstan). The local transmission and persistence of genomes A4, A2a and A3 was also observed in the studied locations. The most prevalent genomes with patterns of variance (confined in a cluster) remain unclassified, and are here proposed as A4-clade based on its divergence within the A cluster.

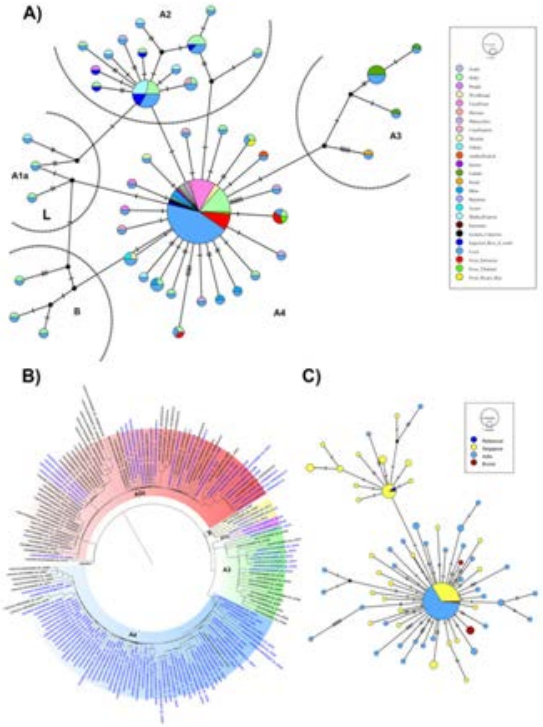


Figure: **Haplotype network and phylogenetic analysis of SARS-CoV-2 sequences.**