Abstract ID: - 92

Abstract Topic: - Rare disease therapeutics

Abstract Title: - Inferring Variants of Uncertain Significance (VUS) in Rare Disease Genetics: An India-centric Study

Presenting author name: - Anjali Krishna A

Presenting author institute: - Amrita School of Biotechnology, Amrita Vishwa Vidyapeetham, Clappana PO 690525, kerala, India

Co-authors name: - Praveen Mathur, Krishna Mohan Medicherla, Renuka Suravajhala, Bipin G Nair, Prashanth Suravajhala

Co-authors institute: - Department of Pediatrics, SMS Medical College and Hospitals, JK Lon, Jaipur 302004, Rajastha, India, Department of Biotechnology and Bioinformatics, Birla Institute of Scientific Research, Jaipur 302001, Amrita School of Biotechnology, Amrita Vishwa Vidyapeetham, Clappana PO 690525, kerala, India, Amrita School of Biotechnology, Amrita Vishwa Vidyapeetham, Clappana PO 690525, kerala, India, Amrita School of Biotechnology, Amrita Vishwa Vidyapeetham, Clappana PO 690525, kerala, India

Aims: - Genetic variation was believed to be associated with exonic regions only. While much emphasis has been placed on coding regions of the genome, non-coding RNAs (ncRNAs) have emerged as essential players in gene regulation and disease mechanisms. Furthermore, genetic variation is beginning to be understood in ncRNAs. However, not much studies have taken place on genetic variation attributing to ncRNAs in rare diseases. There are approximately 6000 rare diseases in the world and over 1000 are reported in India. Many missense and nonsense mutations are known, and variants of uncertain/unknown significance (VUS) are emerging to be understood for causing the rare diseases. In this work, we exploit meta-analyses and aim to identify genetic variation attributing to rare diseases of the Asia/world and India. We also contemplate identifying variants specific to Congenital Pouch Colon (CPC), an anorectal malformation (ARM) that was largely studied and further identify variants specific to syndromic/nonsyndromic ARM from samples sequenced and analyzed in our lab.

Methods: - The genetic diversity inherent in human populations plays a pivotal role in health and disease susceptibility. This study embarks on a comprehensive exploration of genetic variants within three overarching population categories: Indian, Asian, and worldwide. The initial step involved data acquisition from these databases, ensuring that the selected datasets provide an accurate representation of genetic variation within each population category. The data were collected from well-established genetic databases, ensuring inclusion of population-specific information by employing Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. Filters were applied to segregate variants present in the Indian, Asian, and worldwide populations.To this end, we harness the data from two prominent sources, the "indiGenomic VCF(Indian Genomic DB)" and "NCBI Rare Disease" databases, to illuminate the genetic tapestry that unites individuals across diverse backgrounds. Functional impact assessment was then employed to variants with potential biological significance and subsequently

variants were analyzed based on their minor allele frequency (MAF), with thresholds set to capture common, rare, or population-specific genetic diversity. Population-specific MAF data were utilized whenever available, offering insights into genetic variation tailored to each population.

Results: - We delved upon deciphering the candidate genes and their variants from the Asian/world and Indian sub-population and mapped the non-coding variants precisely. The common variants between Indigenomes and our annotation were searched and we found that there are 155 common variants with 15 of them attributing to ncRNA variants, and 3 of those to be pathogenic in Asian/world datasets. Whereas, 64 were commonly found variants with 5 of them attributing to ncRNA and 2 of those classified as pathogenic. We also sought to check the common variants or similar alleles attributing to rare diseases from our cohort. The results of this study present a rich and diverse landscape of genetic variants within these three population categories. The variants were visualized and interpreted with an emphasis on functional impact and allele frequency distribution.

Conclusions: - In conclusion, we explored common genetic variants across Indian, Asian/ worldwide populations underscoring the need for finding candidate variants attributing to pathogenesis and ncRNAs. We believe an investigation into shared genetic diversity offers an insightful glimpse into the human genetic commonality that transcends geographical boundaries and ancestry.

Keywords: - We delved upon deciphering the candidate genes and their variants from the Asian/world and Indian sub-population and mapped the non-coding variants precisely. The common variants between Indigenomes and our annotation were searched and we found that there are 155 common variants with 15 of them attributing to ncRNA variants, and 3 of those to be pathogenic in Asian/world datasets. Whereas, 64 were commonly found variants with 5 of them attributing to ncRNA and 2 of those classified as pathogenic. We also sought to check the common variants or similar alleles attributing to rare diseases from our cohort. The results of this study present a rich and diverse landscape of genetic variants within these three population categories. The variants were visualized and interpreted with an emphasis on functional impact and allele frequency distribution.