

COMMENT



On Clustering of Juvenile Canavan disease in an Indian community due to population bottleneck and isolation: Genomic signatures of a founder event

Ranajit Das ¹✉

© The Author(s), under exclusive licence to European Society of Human Genetics 2022

European Journal of Human Genetics (2023) 31:7–8; <https://doi.org/10.1038/s41431-022-01221-8>

Canavan disease (CD) is an autosomal recessive neurodegenerative disorder with a prevalence of 1 in 200,000–400,000 individuals. While CD occurs among individuals from various ancestral groups, the highest prevalence has been reported among the Ashkenazi Jewish (AJ) people, with a carrier frequency of 1:82 [1]. The Mild/juvenile form (M/JCD) is less common and children with M/JCD may experience mildly delayed speech or motor development. The article by Kotambail et al. titled “Clustering of Juvenile Canavan disease in an Indian community due to population bottleneck and isolation: Genomic signatures of a founder event” reported a novel pathogenic variant *NM_000049.4(ASPA):c.526G>A*, associated with M/JCD in two independent families from Southern India, belonging to the Telugu Devanga Chettiar (TDC) community. They further investigated the possibility of a founder event and subsequent century-old practice of marriage within community (endogamy) leading to the high prevalence of M/JCD among TDCs.

The authors employed various in silico tools to characterize the novel homozygous variant c.526G>A (NC_000017.10:g.3386886G>A) present in the third exon of *ASPA* on chromosome 17 (chr17:3483592 (GRCh38)) and predicted a splicing defect wherein the consensus splice donor site is likely replaced by a cryptic splice donor site 4 bp upstream. While investigating a potential founder event, they found higher fixation index (F) and lower heterozygosity among TDC compared to other individuals with no known history of consanguineous marriage, residing in the same geographical area. The authors further reported high autozygosity i.e. the two homozygous alleles are inherited from a common ancestor (identical by descent (IBD)) among TDC. The fraction of the genome that is IBD among TDC founders was further discerned to be higher than in pairs of unrelated individuals. High autozygosity and IBD sharing observed among the TDC suggest a genetic drift event caused by either endogamy or population bottleneck or even both, as aptly interpreted by the authors. Principal Component Analysis (PCA) revealed distinct clusters of TDC and Tamil-speaking population (TML) from the same geographical area. While the differentiation between TDC and TML is not startling, the clustering of ITU and GIH from 1000 Genomes project is surprising, as the ITU has been shown to cluster mostly with the South Indian populations and are genetically mostly distinct from GIH [2]. Further, the PCA plot and the maximum likelihood (ML) tree are not in complete agreement. If

we go by the ML tree, GIH should cluster with NOI (North Indians) and ITU should be in closer proximity to TDC and TML, which is supported by the previous studies, but is not reflected by the PCA plot in this study. This disagreement could be attributed to the smaller sample size of TDC, TML and NOI utilized here. A potential way to address this might be to include more publicly available South Asian genomes to resolve the ML tree and identify finer structures within the TDC genomes. Overall, the population genetic analyses indicates that after the initial founder event, the TDC genomes have remained in long-term genetic isolation, preserved by endogamy, with little or no gene flow into the community.

Devanga Chettiars are one of the oldest weaving communities in South India. They speak Old-Kannada, Telugu and Tamil. Their origin and migration history has remained elusive. Folklore suggests that they lived near the borders of the Southern Indian states of Tamil Nadu and Karnataka, across the Western Ghats for thousands of years preceding the Vijaynagara Empire and were known as Cheniyars/Chedars at this time. It seems the word ‘Devanga’, which literally means the ‘Body of the God’ came to be used for them later. Notably, both the Southern Indian languages, Telugu and Kannada have evolved from Old-Kannada, arguably the spoken language of Devanga Chettiars’ ancestors. Till date, many Devangas speak Old-Kannada without much influence of Sanskrit and Prakrit. For centuries, Devanga Chettiars have been an actively migrating community mostly due to business and trading. It may be speculated that during the illustrious Vijaynagara Empire (fifteenth–sixteenth century AD), they had migrated to various parts of the neighboring Southern Indian states of Andhra Pradesh and settled there [3]. It is plausible that once in Andhra Pradesh, they began speaking Telugu, the local language which may have been facilitated because of the similarities between Old-Kannada and Telugu languages. Notably, Telugu and Kannada evolved into their own around thirteenth century AD, only ~ 300 years before Devangas’ migration to Andhra. Overall, if local folktales are to be believed, Devangas appear to have a deep cultural root in Tamil Nadu but are genetically more proximal to ancient Kannadiga people.

To investigate the genetic similarities of TDC with modern-day indigenous tribal populations and other mainstream populations from Karnataka, Andhra Pradesh, Telangana, and Tamil Nadu, five TDC genomes were merged with 99 populations across Karnataka,

¹Yenepoya Research Centre, Yenepoya (Deemed to be University), University Road, Deralakatte, Mangalore 575018 Karnataka, India. ✉email: das.ranjit@gmail.com

Received: 10 October 2022 Accepted: 14 October 2022

Published online: 1 November 2022

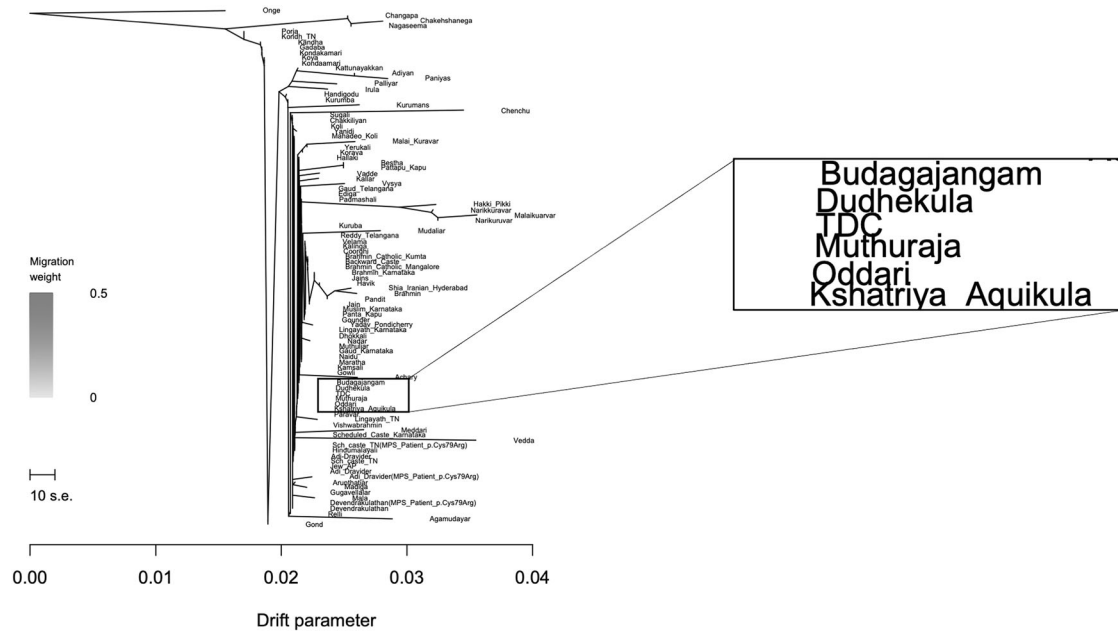


Fig. 1 Maximum Likelihood (ML) Tree of various Southern Indian populations. The ML was constructed using TreeMix v1.13. Individuals across states of Karnataka, Tamil Nadu, Andhra Pradesh and Telangana were employed in the ML tree.

Tamil Nadu, Andhra Pradesh, and Telangana. A maximum likelihood (ML) based tree was constructed using TreeMix v1.13 (Fig. 1).

The ML tree demonstrated the genetic proximity of TDC with the Telugu-speaking Dudekula people, and Tamil and Telugu-speaking Muthuraja people, whose origin has also remained elusive till date. TDC genomes also revealed genetic similarities with Budaga Janam people from Karnataka, Andhra Pradesh and Tamil Nadu, and Telugu-speaking Oddari people. Overall, the ML tree revealed that TDC had socio-cultural connection with the native Telugu speakers from Andhra Pradesh, having adopted the local language Telugu after their migration to Andhra Pradesh.

One may surmise that around 400 years ago, likely after the fall of the Vijayanagara Empire in the sixteenth century AD, the gene flow between the TDC and populations across Andhra and Tamil Nadu ended or was severely curtailed. The novel mutation described via mutation dating in Kotambail et al. likely appeared during or soon after the genetic isolation of TDC. Due to the practice of endogamy for last several hundred years, this mutation has been preserved in the TDC community. Interestingly, similar to TDC, AJs, among whom the highest prevalence of CD has been reported, practice ethnic endogamy to preserve their ancestral gene pool and thus remain genetically isolated from their European neighbors. It can be speculated that in case of both TDC and AJs the causative mutations of CD are maintained by endogamy.

Overall, the article by Kotambail et al. utilizes medical and population genetics approaches to describe a novel mutation leading to M/JCD in two families that belong to an endogamous Southern Indian population. The confluence of these two streams of investigation in human genetics is not frequent in published literature from India. Given that India is a melting pot of genetic heterogeneity and potpourri of socio-cultural diversity, the ancestry and migration history of various communities and the origin of numerous associated novel mutations remain unknown. More studies such as this with the collaboration of population and medical genetics strategies will not only improve our understanding of rare disorders to accelerate the development of individual and population-specific therapeutics but further

provide more granular insight into human history from the Indian subcontinent.

REFERENCES

1. Fares F, Badarnah K, Abosaleh M, Harari-Shaham A, Diukman R, David M. Carrier frequency of autosomal-recessive disorders in the Ashkenazi Jewish population: should the rationale for mutation choice for screening be reevaluated? *Prenat Diagn.* 2008;28:236–41.
2. Das R, Upadhyai P. Application of geographic population structure (GPS) algorithm for biogeographical analyses of populations with complex ancestries: a case study of South Asians from 1000 genomes project. *BMC Genet.* 2017;18:109.
3. Ramaswamy V. Mapping migrations of South Indian weavers before, during and after the Vijayanagar period: thirteenth to eighteenth centuries. *Globalising Migration History: Brill;* 2014. p. 89–121.

ACKNOWLEDGEMENTS

I sincerely acknowledge Dr. Gautham Arunachal, Dr. Ananthapadmanabha K and team from National Institute of Mental Health and Neuro Sciences (NIMHANS), Bengaluru, Karnataka, India for kindly providing me the genome-wide genotype data of the TDC individuals.

FUNDING

There was no specific funding available for this work.

COMPETING INTERESTS

The author declares no competing interests.

ADDITIONAL INFORMATION

Correspondence and requests for materials should be addressed to Ranajit Das.

Reprints and permission information is available at <http://www.nature.com/reprints>

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