Recurrence of reported GNPTAB, GNPTG and NAGPA gene variations associated with stuttering - an evaluation

SRIKUMARI CR[1], NANDHINI G[1], Jeffrey JM[1], and MATHURAVALLI KRISHNAMOORTHY[1]

[1]University of Madras

October 8, 2020

Abstract

Stuttering is a childhood onset fluency disorder, intertwined with physiological, emotional and anxiety factors. The present study was designed to evaluate the recurrence of the reported mutations among three previously implicated (GNPTAB, GNPTG, NAGPA) candidate genes, in persons with stuttering (PWS) from south India. Mutation screening was performed on 64 probands on 12 specific exons, by Sanger sequencing. A total of 12 variants were identified, which included five nonsynonymous missense, five synonymous and two non coding variants. Only three unrelated probands, harbored heterozygous likely pathogenic missense variants (c.3598G>A in GNPTAB, c.802A>C in GNPTG and c.131G>C in NAGPA) resulting in an overall frequency of 4.7% and an allele frequency of 2.3% (3/128*100). Among the three likely pathogenic variants only two co-segregated (c.3598G>A in GNPTAB - STU 29 and c.802A>C in GNPTG - STU 63) with the affected status reducing the likely pathogenic allele frequency to 1.6% (2/128*100). The recurrence of pathogenic variants in our study corroborate the causative role of these genes in stuttering but still remains unknown as to how the speech dysfluency occurs even in its heterozygous condition.
Keywords: Stuttering, candidate genes, GNPTAB, GNPTG, NAGPA

INTRODUCTION

Speech is a unique motor function and when affected causes both receptive and expressive communication disorders, significantly reducing the quality of life. Stuttering is an expressive fluency disorder, characterized by repetitions, prolongations, blocks, along with secondary behaviors (head jerks, lip tremors and eye blinks) and often lead to psychological problems such as increasing anxiety[1].

Developmental stuttering arises in children of 2-5 years age group, but most of them (80%) recover spontaneously. But there is greater chance of recovery (male female ratio ˜5:1) among females[2–4]. Prevalence of stuttering ranges from 0.3% to 5.6% and the average prevalence over the lifespan may be lower than the commonly cited 1%[5]. In a recent study based on 75000 school children in India, we reported a prevalence of 0.46%[6].

Research studies on the etiology of stuttering focused mainly on neuroimaging and genetics. Most of brain imaging methods have consistently reported structural or functional differences contributing to inefficient communication in stuttering. It includes over activity of the dopamine neurotransmitter[7], abnormal functional lateralization of cortical connections[8], deficits in white matter tract that connects motor and auditory structures, corpus callosum as well as cortical and subcortical areas[9]. There is also growing consensus about the genetic origin of Central Nervous System dysfunctions [10].

Genetic dissection is challenging due to gene-gene/gene-environment interactions, genetic heterogeneity, gender bias, incomplete penetrance and phenocopies [11]. Initial linkage studies found suggestive evidence for chromosome regions (1, 2q, 3q, 5q, 7q, 9p, 9q, 13q, 15q, 18p, 18q, 20p) implicated in stuttering but with little overlap across studies [12,13,14]. However definitive evidence for linkage was identified on chromosomes

3, 12 and 16 in highly consanguineous Pakistani families[15,16,17] and on chromosomes 2, 3, 14 and 15 in a large Cameroon family[18]. Although linkage studies are spread across Hutterite, European and American population, the four genes,*GNPTAB, GNPTG* , *NAGPA* and *AP4E1,* identified are restricted to two regions [Pakistan[19]; Cameroon[20]] with distinct ethnicities. The combined contribution of these genes were estimated to be 20%[21]. All these genes point to intracellular trafficking deficits[22]. Indeed genomics of stuttering is an emerging field where the involvement of new genes are yet to be identified.

GWAS study suggested ten candidate genes (*FADS2, PLXNA4, CTNNA3, ARNT2, EYA2, PCSK5, SLC24A3, FMN1, ADARB2* and non-coding RNA*RNU6-259P*) involved in neural pathways[23]. Mutations so far identified implicate lysosomal dysfunction but the biological mechanisms that affect speech are under investigation. Mice models also point to deficits in inter-hemispheric connectivity in astrocytes of corpus callosum [24,25], thus linking genes to brain activity. Stuttering genes *GNPTAB* and *GNPTG*were also previously implicated in a rare lysosomal storage disorder - mucolipidosis.

From the genetic perspective, genes identified play role in targeting enzymes to lysosomes that is crucial for biogenesis and also in the maintenance of myelin sheaths. From neurological perspective hyperactivity of dopamine and the white matter abnormalities observed in stuttering, provide a possible neurochemical basis but the effect of the mutations in neural cell biology is still unexplored. Owing to significant plasticity of brain it was unable to account for the observed differences among PWS and control, as to whether they are cause or result of stuttering[26]. Thus the connecting dots between dopamine, neural circuits and cellular waste disposal is yet to be connected.

No studies from India are available till date that implicate any genes for stuttering. This gap motivated us to first ideally ensure the frequency of the previously implicated genes for stuttering in our population, before initiating advanced approaches. We evaluate the recurrence of the reported mutations among the three reported (*GNPTAB, GNPTG, NAGPA)* stuttering candidate genes in PWS from south India. Attempts to employ identical experimental design to concurrently verify and replicate the findings independently would on one hand help understand ethnicity specific variations and on another hand would also enable reproducibility of results and facilitates pooling of data during meta-analysis.

RESULTS

Mutational analysis of the three putative genes in stuttering

A total of 64 unrelated probands with non-syndromic persistent stuttering (sex ratio - 12:1; 59:5; mean age at onset of 5.13 years) were screened for the recurrence of mutations in the three stuttering implicated genes. Sixty seven percent (43/64) of them had family history. More than 50% of PWS were found to be severe; **53.1% severe** (34/64), **28.1% moderate** (18/64) and **18.8% mild** (12/64).

Mutation screening of the twelve specific exons previously reported (figure 1*),* identified a total of **12 variants** that includes five nonsynonymous missense variants, five synonymous and two non coding variants (tables 1 and 2; figures A1-A3). The distribution of these 12 variants among the 64 probands are shown in figure 2a. Variants observed in *NAGPA* (n=6) were higher than that of *GNPTAB*(n=2) and *GNPTG* (n=4) (figure 2b).

Only three unrelated probands (STU 29, STU 63 and STU 34), harbored heterozygous likely pathogenic missense variants (c.3598G>A in *GNPTAB* , c.802A>C in *GNPTG* and c.131G>C in *NAGPA* respectively) with an overall frequency of 4.7% (3/64*100) and an allele frequency of 2.3% (3/128 * 100). None of the three showed more than one pathogenic allele but there was co-occurrence of synonymous and non-coding variants (table 3).

The two missense variants (c.139C>T & c.1394 C>T) in *NAGPA,* had low conservation scores and were found in high frequency in the ExAC database supporting their benign nature. Hence, segregation analysis and genotype-phenotype correlations were analyzed only for the three likely pathogenic variants (figure 3-5). Only two variants (c.3598G>A and c.802A>C) co-segregated (table 4) with the affected status, reducing the likely pathogenic allele frequency to 1.6% (2/128 *100).

Impact of a *de novo* variant (c.802A>C/+ in*GNPTG*) identified in STU 63 family

To study the impact of a *de novo* heterozygous variant (c.802A>C/+) in *GNPTG* gene identified in one family (STU 63), mRNA expression profile and lysosomal enzyme study was performed along with mucolipidosis screening test. All the members of the family including the affected proband and unaffected father, mother and sister screened, were found negative for mucolipidosis test. The activity of the enzymes studied in plasma was found to be well within the normal range (table 5).

To quantify mRNA, the CT data obtained was used to calculate ΔCT values (ΔCT = CT target – CT reference) that are normalized to the housekeeping β-actin gene for each of the target gene studied (*GNPTAB* , *GNPTG* and *NAGPA*) and plotted in the figure 6. The data suggests that there is variability within the controls (father, mother and sister) and there is no obvious difference between the proband and the internal control group.

The 2ˆ- ΔΔCT calculations were performed to check if the control data can be clubbed. The ΔΔCT value was obtained by subtracting the ΔCT of proband with ΔCT of control (ΔΔCT = ΔCT test sample – ΔCT control). Because of the variability, and smaller sample size, 2- ΔΔCT calculation and statistical analysis was not possible. Also averaging the control ΔCt values and comparing it with the respective gene expression value in the single proband would be misleading. However, the data suggests that there is no apparent differences in the expression of *GNPTG* ,*GNPTAB* and *NAGPA* genes between the proband and the controls.

DISCUSSION

Speech is a robust faculty that serves most people in the face of various challenges. Nevertheless, developmental stuttering is a complex phenotype with overt diversity in terms of both genetic and deterministic risk factors.

Mutational analysis of implicated genes in stuttering

To the best of our knowledge this is the first study to investigate three functionally related genes *viz.,* *GNPTAB* , *GNPTG* and*NAGPA* implicated in stuttering, from India. We focused on the recurrence of the previously reported mutations in 64 probands. In a total of **12 variants** identified only two co-segregated (c.3598G>A in *GNPTAB* - STU 29 and c.802A>C in *GNPTG* - STU 63) with the affected status resulting in a likely pathogenic allele frequency of 1.6%.

A meta-analysis of worldwide unrelated PWS, identified 81 rare nonsynonymous coding variants, in either of the three putative genes, accounting for a frequency of 16% (164/1013)[20]. In our study, we observed four rare nonsynonymous coding variants accounting for 6% (4/64). Among the twelve variants identified in this study, five of them (c.3598G>A, c.1932A>G in*GNPTAB* and c.131G>C, c.333 A>G, c. 1485C>T in *NAGPA*) were previously reported in population with stuttering. One variant (c.702T>C in*GNPTG*) was reported in mucolipidosis III. However, the remaining six variants (c.802A>C, c.813G>A, -4 C>T in *GNPTG* and c.139C>T, c.1394 C>T, c.1174+53C>A in *NAGPA*) were just reported in the ExAC database. The three likely pathogenic nonsynonymous variants are discussed briefly:

(i) STU 29 family with c.3598G>A variant in*GNPTAB* gene

The fact that the highest linkage scores were obtained for this variant combined with lack of other plausible genetic variant within the linkage interval suggested increase risk of stuttering when present in either one or two copies [19]. Fedyna *et al.,* [27] reported 4/8 unrelated PWS carried atleast one copy of p.Glu1200Lys mutation in *GNPTAB* gene and established this as a founder mutation in Asian population, originating from Pakistan or India. Recurrence of this lysine variant in heterozygous condition (0.8%) that is segregating with affected status among our south Indian stuttering family **favors the founder effect** in Asians. Stuttering endophenotypes are homogenous and stable phenotypes. We hypothesized that in a cohort with severe stuttering as an endophenotype, there may be an increased chance to identify lysine variants in homozygous condition. In order to verify this we tested additionally, 26 severe PWS, but identified again only heterozygous

lysine variants in three of them; thereby increasing the overall frequency of this variant to 4.4% (4/90*100) with an allele frequency of 2.2% (4/180*100).

High frequency of this variant among south Asian ancestry (2.1%) in the ExAC database, questions its pathogenicity. However, it is highly conserved (Consurf = 8) and a mutation in it disrupts the helical segment that may be crucial for interaction with other subunits or proteins in lysosomal pathway (figure A4).

In a recent animal model study 3- to 8-day old mice pups were engineered to carry two copies of the lysine mutation ($Gnptab^{mut/mut}$), resulting in significantly longer pauses in their spontaneous vocalizations consistent with some features of human stuttering, but was not found to be lethal as in case of mutations reported in mucolipidosis. This was neither observed in littermates without the mutation ($Gnptab^{wt/wt}$) nor in heterozygous ($Gnptab^{mut/wt}$) littermates[24]. Though the causative role is well established for the homozygous lysine variant in *GNPTAB* gene, the heterozygous variants were similar to wild-type phenotypically in mouse model. However it is still unclear how a recessive allele in heterozygous condition is causing stuttering phenotype in human model; this trend has been consistently observed not only in our study but in all other reported studies[19,21,28]. This may be due to (i) a second pathogenic mutation in this gene acting *in trans* or (ii) other additive genetic factors that may play a causative role or (iii) the gene may have a role in some other unknown pathway. Hence identifying new interacting genes or pathways may clarify the causative role of this gene in stuttering. In fact, dominance and recessiveness are not essentially allelic properties but measured in relation to the effects of other alleles at the same locus. Additionally, dominance may change according to the level of organization of the phenotype and its variations highlight the complexity of understanding genetic influences on phenotypes[29].

Two more *GNPTAB* homozygous mutations p.Ser321Gly and p.Ala455Ser were engineered in mice, that also displayed vocalization deficits traceable to abnormalities in astrocytes of corpus callosum[25].

(ii) STU 63 family with c.802A>C variant in *GNPTG* gene

Since this variant was found only in the proband, it may be a *de novo* variant. This heterozygous variant was not observed so far in stuttering population, but reported in ExAC database.

In general, *de novo* being a rare genetic variant, may be more deleterious than an inherited variant since they are less subjected to evolutionary selection. They may be prime candidates when genetic diseases occur sporadically[30]. In our study the role of this *de novo* variation in stuttering is ostensibly supported by high conservation score and its absence in unaffected. Recurrence of this mutation in unrelated PWS may provide further evidence for its role.

Since most of the mutations so far reported in stuttering are by and large heterozygous, similar to that observed in our study, we wanted to comprehend how heterozygous mutations are involved in stuttering. Hence to study the impact of the *de novo* heterozygous missense variant identified in *GNPTG* , (i) quantification of mRNA by RT-PCR (ii) activity of lysosomal enzymes in plasma was carried out. We assessed if there were any differences in the expression of targeting genes and also targeting function of the lysosomal enzymes between the affected and unaffected members.

If the variation affects the targeting function, the enzyme will not be targeted to lysosomes but will be secreted in plasma. Thus the enzyme deficiency can be demonstrated by elevated enzyme activity in plasma[31]. Nevertheless, in our study the activity of lysosomal enzymes were not elevated in plasma, indicating that the enzyme might be successfully targeted to lysosomes. We propose that, since the variation observed is in heterozygous condition, either the normal copy is sufficient or this variation does not affect the function of the enzyme. Similarly, there was no fold change in the mRNA level of the three genes between the affected (proband) and unaffected members (father, mother and sister) of the family. Hence it was difficult to conclusively demonstrate the pathogenicity of this *de novo*mutation in stuttering.

(iii) STU 34 family with c.131G>C variant in *NAGPA* gene

We observed this variant in both affected and unaffected family members. This may be explained by incomplete penetrance that may fail to show any symptoms in unaffected or could be due to phenocopies in affected members who may not be real carriers of variant but tend to display stuttering under environmental effects[19]. Also variants observed in normal individuals may cause stuttering but left un-informative owing to early recovery[28] but in our family no such recovery was reported.

Since its frequency was also low in ExAC database the role of this variant remains inconclusive. In addition to this the conservation score is also found to be low across the species. Only one study[21] has reported this variant among stuttering population of European descent.

Role of synonymous and noncoding variations

Overall five synonymous variants and two noncoding variants were observed in our cohort. Synonymous mutations are often considered as silent mutations due to degeneracy of genetic code. But they may have important consequences and is now recognized to be crucial in influencing gene expression, conformation and in cellular function[32]. Although our study shows only three probands to harbour likely pathogenic alleles there is a preponderance of synonymous and noncoding variants among all the stuttering individuals screened. Complex disorders often tend to have multiple mutations. A mutation may not be detrimental individually but the joint effect of multiple variants in the same gene or different genes can contribute to a disorder but however predictions are limited to single variant[33].

In our study the recurrence of the pathogenic variants in lysosomal pathway corroborates the causative role for them in stuttering. The importance of the implicated genes can be understood by its recurrence in other ethnic populations. Identification of recurrent mutations helps in cost effective screening in a large sample of PWS. However, it should be borne in mind that scope of this study is limited because our screening includes only the exons that were previously reported.

CONCLUSION

Decades of research has finally tied stuttering to certain genes and changes in brain. Mutation screening of the three implicated genes (*GNPTAB* , *GNPTG* and *NAGPA)* among 64 PWS, resulted in a likely pathogenic allele frequency of 1.6%. Recurrence of mutations in the three genes among our south Indian stuttering cohort corroborates the causative of these genes to stuttering. Thus mutational screening ended up with a minimal resolution of 3.1% (2/64) that could be ascribed to these genes but remains inconclusive. Hence involvement of more stuttering genes are predicted and can certainly be addressed using next generation sequencing technology. Since stuttering is a complex disorder two highly multiplex families were chosen from existing database[6] to identify new genes involved in related pathways using exome sequencing in the second paper submitted in the series.

METHODOLOGY

This study was approved by Institutional Ethical Committee and informed consent was obtained from all participants. Sixty four probands were recruited from various schools, hospitals and speech therapy clinics and clinically diagnosed for stuttering by speech pathologist using Stuttering Severity Instrument 3(SSI-3)[34]. A structured interview using questionnaire was conducted to elicit demogenetic details (table A1). More details on recruitment of the probands is given in our previous paper[6].

Eight milliliters of blood was collected by venipuncture into labelled EDTA coated vacutainers (Beckon and Dickinson Co., USA). Genomic DNA was isolated using Phenol-Chloroform extraction method[35].

Mutational analysis of *GNPTAB, GNPTG* and *NAGPA*gene variants

The 12 specific exons spanning across the three genes viz.,*GNPTAB* , *GNPTG* and *NAGPA* implicated in stuttering were screened (Figure 1). Primer sequences were adapted from Kang*et al.,* (2010), after improvising (NCBI's Primer-BLAST) the sequence coverage of exon 10 of *NAGPA* gene.

The amplified PCR products were purified by the FavorPrep^{TM} PCR purification kit (FAVOURGEN, Taiwan). The amplicons were sequenced using ABI Prism Big-Dye Terminator 3.1 cycle sequence reaction kit on ABI 3730XL automated sequencer (Applied Biosystems, USA). Chromatograms were analyzed using NCBI nucleotide BLAST (www.ncbi.nlm.nih.gov/) and UCSC genome browser BLAT (genome.ucsc.edu).

Variants identified were predicted using VarSome (https://varsome.com/; tools include various predictors like DANN, Mutation taster, Likelihood Ratio Test - LRT, Mutation assessor, SIFT, Provean etc) and Polyphen tool, to deduce the pathogenicity. Cosegregation of the pathogenic variants among the family members was also evaluated. Novelty and frequency of the variations were compared with ExAC database.

To study the impact of a *de novo* heterozygous variant**(c.802A>C/+)** in *GNPTG* gene identified in one family (STU 63), mRNA expression profile and lysosomal enzyme study was performed along with mucolipidosis screening test. Since the proband alone was affected, while his father, mother and sister served as controls. The plasma collected from fresh blood (5 ml) was used to study the enzyme activity. RNA was isolated using mirVana miRNA isolation kit (Invitrogen, USA) as per manufacturer's instructions and checked for integrity and purity.

A two step qRT-PCR was used to measure the transcript levels of the mRNAs of interest.

1. From 500ng of total RNA, **cDNA was synthesized** using the RevertAid First Strand cDNA Synthesis Kit (Thermo Scientific, USA) according to manufacturer's instruction in ABI GeneAmp 9700 PCR System. Reverse transcription reaction mix of 20µl was prepared and was loaded on to ABI GeneAmp 9700 PCR System.
2. For **quantifying gene expression** real-time Quantitative PCR was performed on QuantStudio3 Real-Time PCR System using GoTaq DNA polymerase (Promega) in the presence of SYBR Green. The primers specific for the transcripts of GNPTAB, GNPTG and NAGPA (table A2) were designed with the aid of IDT software and checked for specificity with NCBI's primer-blast. The annealing temperature of the primers were optimized using temperature gradient PCR. Reaction mix for the samples under investigation along with NTC (no template control) was prepared in triplicates for *GNPTAB* , *GNPTG* , *NAGPA*and *β-actin* genes. The reaction mix was loaded on to a 96 well plate and sealed with MicroAmp® Optical Adhesive Film (Applied Biosystems). Ct value (cycle threshold) is the number of PCR cycles required to achieve a given level of fluorescence. Since Ct value is proportional to logarithm of initial amount of the target, the relative concentration of one target with another is reflected as a difference in cycle number (ΔCt) that is necessary to achieve equivalent level of fluorescence. The expression levels of GNPTAB, GNPTG and NAGPA were measured by relative quantification using the ΔCT method with β-actin as endogenous control.**Statistical analysis:** Delta Ct values were normalized to the housekeeping β-actin gene for each of the target gene (*GNPTAB* ,*GNPTG* and *NAGPA*). Statistical analysis was performed by averaging the control Delta Ct values and comparing it with the respective gene expression. Lysosomal targeting of proteins was studied using specific substrate for Arylsulphatase A, Hexosaminidase A and β galactosidase enzymes, with plasma samples from stuttering proband and controls. All members in the family were also evaluated for ML phenotype using a rapid calorimetric screening method.

Bibliography

1. Ashurst J, Wasson M. Developmental and persistent developmental stuttering: An overview for primary care physicians. J Am Osteopath Assoc. 2011;111(10):576–80.

2. Drayna D, Kilshaw J, Kelly J. The sex ratio in familial persistent stuttering. Am J Hum Genet. 1999 Nov;65(5):1473–5.

3. Yairi E, Ambrose NG. Early childhood stuttering I: persistency and recovery rates. J Speech Lang Hear Res. 1999 Oct;42(5):1097–112.

4. Yairi E, Ambrose N. Early childhood stuttering. Austin, TX: Pro-Ed. Inc; 2005.

5. Yairi E, Ambrose N. Epidemiology of stuttering: 21st century advances. Journal of fluency disorders. 2013 Jun 1;38(2):66-87.

6. Nandhini Devi G, Thalamuthu A, Valarmathi S, Karthikeyen NP, Srikumari Srisailapathy CR. Genetic epidemiology of stuttering among school children in the state of Tamil Nadu, India. J Fluency Disord. 2018;58.

7. Cykowski, M. D., Kochunov, P. V., Ingham, R. J., Ingham, J. C., Mangin, J. F., Riviere, D., Lancaster, J. L. &Fox PT. Perisylvian sulcal morphology and cerebral asymmetry patterns in adults who stutter. Cereb Cortex. 2008;18(3):571–83.

8. Chang S-E. Research Updates in Neuroimaging Studies of Children Who Stutter. Semin Speech Lang,. 2014;35(2):67–79.

9. Chang, S. E., Zhu, D. C., Choo, A. L., & Angstadt M. White matter neuroanatomical differences in young children who stutter. Brain. 2015;138(3):694–711.

10. Ingham RJ. Brain imaging studies of developmental stuttering. J Commun Disord. 2001;34(6):493–516.

11. Oliveira, B. V. D., Domingues, C. E. F., Juste, F. S., Andrade, C. R. F. D., & Moretti-Ferreira D. Familial persistent developmental stuttering: genetic perspectives. Rev da Soc Bras Fonoaudiol. 2012;17(4):489–94.

12. Shugart YY, Mundorff J, Kilshaw J, Doheny K, Doan B, Wanyee J, et al. Results of a genome-wide linkage scan for stuttering. Am J Med Genet [Internet]. 2004 Jan 15;124A(2):133–5. Available from: http://doi.wiley.com/10.1002/ajmg.a.20347

13. Suresh R, Ambrose N, Roe C, Pluzhnikov A, Wittke-Thompson JK, Ng MC-Y, et al. New Complexities in the Genetics of Stuttering: Significant Sex-Specific Linkage Signals. Am J Hum Genet. 2006;78(4):554–63.

14. Wittke-Thompson, J. K., Ambrose, N., Yairi, E., Roe, C., Cook, E. H., Ober, C., & Cox NJ. Genetic studies of stuttering in a founder population. J Fluen Disord. 2007;32(1):33-50.

15. Raza MH, Riazuddin S, Drayna D. Identification of an autosomal recessive stuttering locus on chromosome 3q13.2–3q13.33. Hum Genet. 2010 Oct 13;128(4):461–3.

16. Riaz N, Steinberg S, Ahmad J, Pluzhnikov A, Riazuddin S, Cox NJ, et al. Genomewide significant linkage to stuttering on chromosome 12. Am J Hum Genet. 2005;76(4):647–51.

17. Raza MH, Amjad R, Riazuddin S, Drayna D. Studies in a consanguineous family reveal a novel locus for stuttering on chromosome 16q. Hum Genet. 2012 Feb;131(2):311–3.

18. Raza MH, Gertz EM, Mundorff J, Lukong J, Kuster J, Schäffer AA, et al. Linkage analysis of a large African family segregating stuttering suggests polygenic inheritance. Hum Genet. 2013;132(4):385–96.

19. Kang C, Riazuddin S, Mundorff J, Krasnewich D, Friedman P, Mullikin JC, et al. Mutations in the lysosomal enzyme-targeting pathway and persistent stuttering. N Engl J Med. 2010 Feb 25;362(8):677–85.

20. Raza MH, Mattera R, Morell R, Sainz E, Rahn R, Gutierrez J, et al. Association between Rare Variants in AP4E1, a Component of Intracellular Trafficking, and Persistent Stuttering. Am J Hum Genet. 2015 Nov 5;97(5):715–25.

21. Raza MH, Domingues CEF, Webster R, Sainz E, Paris E, Rahn R, et al. Mucolipidosis types II and III and non-syndromic stuttering are associated with different variants in the same genes. Eur J Hum Genet. 2016;

22. Frigerio-Domingues C, Drayna D. Genetic contributions to stuttering: the current evidence. Mol Genet Genomic Med. 2017 Mar;5(2):95–102.

23. Kraft SJ. Genome-wide association study of persistent developmental stuttering. University of Illinois; 2010.

24. Barnes, T. D., Wozniak, D. F., Gutierrez, J., Han, T. U., Drayna, D., & Holy TE. A mutation associated with stuttering alters mouse pup ultrasonic vocalizations. Curr Biol. 2016;26(8):1009–18.

25. Han, T. U., Root, J., Reyes, L. D., Huchinson, E. B., du Hoffmann, J., Lee, W. S., . . . & Drayna D. Human GNPTAB stuttering mutations engineered into mice cause vocalization deficits and astrocyte pathology in the corpus callosum. Proc Natl Acad Sci. 2019;116(35):17515–24.

26. Büchel C, Watkins KE. Genetic susceptibility to persistent stuttering. The New England journal of medicine. 2010 Jun 5;362(23).

27. Fedyna A, Drayna D, Kang C. Characterization of a mutation commonly associated with persistent stuttering: Evidence for a founder mutation. J Hum Genet. 2011;56(1):80–2.

28. Kazemi, N., Estiar, M. A., Fazilaty, H., & Sakhinia E. Variants in GNPTAB, GNPTG and NAGPA genes are associated with stutterers. Gene. 2018;647:93–100.

29. Miko I. Genetic dominance: genotype-phenotype relationships. Nat Educ. 2008;1(1):140.

30. Veltman, J. A., & Brunner HG. De novo mutations in human genetic disease. Nat Rev Genet. 2012;13(8):565.

31. Sheth, J. J., Mistri, M., Kamate, M., Vaja, S., & Sheth FJ. Diagnostic strategy for mucolipidosis II/III. Ind Pediatr. 2012;49((12)):975–7.

32. Sauna, Z. E., & Kimchi-Sarfaty C. Understanding the contribution of synonymous mutations to human disease. Nat Rev Genet. 2011;12(10):683.

33. Liu, M., Watson, L. T., & Zhang L. Predicting the combined effect of multiple genetic variants. Hum Genomics. 2015;9(1):18.

34. Riley GD. Stuttering Severity Instrument-3 (SSI-3). 1994;Austin, TX.

35. Sambrook, J.F. and Russell D. Molecular Cloning: A Laboratory Manual. 3rd ed., V. Press CSHL, editor. 2001.

Table 1: Allele frequencies of the 12 variants observed in *GNPTAB, GNPTG* and *NAGPA* genes among the 64 probands with stuttering and their comparison with ExAC database

S.No	GENE	Nucleotide change	Amino acid change	Exon	dbSNP ID	Hom/Het (n=64 probands)	Allele frequency (n=128 alleles)	Stuttering studies	Allele frequency in South Asian ExAC
Missense variants	**Missense variants**	**Missense variants**	**Missense variants**	**Missense variants**	**Missense variants**	**Missense variants**	**Missense variants**	**Missense variants**	**Missense variants**
1	*GNPTAB*	**c.3598G>A**	Glu1200Lys	Exon 19	rs137853825	- /1	0.008	(8/1013cases) = 0.00789	0.02181
2	*GNPTG*	**c.802A>C**	Ile268Leu	Exon 10	rs759796840	- /1	0.008		0.00006064
3	*NAGPA*	**c.131G>C**	Arg44pro	Exon 2	rs374266430	- /1	0.008	(1/1013) = 0.000099	0.000
4	*NAGPA*	c.139C>T	Leu47Phe	Exon 2	rs371054576	- /1	0.008		0.000

S.No	GENE	Nucleotide change	Amino acid change	Exon	dbSNP ID	Hom/Het (n=64 probands)	Allele frequency (n=128 alleles)	Stuttering studies	Allele frequency in South Asian ExAC
5	*NAGPA*	c.1394 C>T	Thr465Ile	Exon 10	rs7188856	- /22	0.172		0.1731
Synonymous variants	**Synonymous variants**	**Synonymous variants**	**Synonymous variants**	**Synonymous variants**	**Synonymous variants**	**Synonymous variants**	**Synonymous variants**	**Synonymous variants**	**Synonymou**
6	*GNPTAB*	**c.1932A>G**	Thr644Thr	Exon13	rs10778148	42/12	0.75	120/1708 alleles	0.6234
7	*GNPTG*	**c.702T>C**	Pro234Pro	Exon 9	rs532275192	- /1	0.008		0.003333
8	*GNPTG*	**c.813G>A**	Thr271Thr	Exon 10	rs377647926	- /18	0.14		0.00004515
9	*NAGPA*	**c.333 A>G**	Gly111Gly	Exon 2	rs2972272	41/19	0.789	229/1708	0.8266
10	*NAGPA*	**c. 1485C>T**	Asn495Asn	Exon 10	rs887854	42/22	0.828		0.8218
Non coding variants	**Non coding variants**	**Non coding variants**	**Non coding variants**	**Non coding variants**	**Non coding variants**	**Non coding variants**	**Non coding variants**	**Non coding variants**	**Non coding variants**
11	*GNPTG*	**-4 C>T**	-	5'UTR	rs554707396	- /1	0.008		0.001656
12	*NAGPA*	**c.1174+53C>A**		intron 7	rs2937112	22/26	0.547		-

Table 2: Pathogenicity prediction of the variants observed in three genes for stuttering using various bioinformatics tools

S.No	Location	Nucleotide change	dbSNP ID	General DANN	General Mutation taster	Functional SIFT	Functional Provean	Conservation LRT	Conservation Mutation Assessor	Poly Phen-2	I Mutant v2.0	C S
Missense variants	**Missense variants**	**Missense variants**	**Missense variants**	**Missense variants**	**Missense variants**	**Missense variants**	**Missense variants**	**Missense variants**	**Missense variants**	**Missense variants**	**Missense variants**	**M v a**
1	*GNPTAB*	**c.3598G>A**	rs137853825	59982	Disease causing	D	Damaging	Deleterious	Low	PD	Decreased stability	9
2	*GNPTG*	**c.802A>G**	rs759796840	09696	Disease causing	D	Neutral	Deleterious	Medium	PD	Decreased stability	8

S.No	Location	Nucleotide change	dbSNP ID	General	General	Functional	Functional	Conservation	Conservation	PolyPhen-2	I Mutant v2.0	C S
3	*NAGPA*	**c.131G>C**	rs374266430	9819	Polymorphism		Damaging	Neutral	Medium	PD	Decreased stability	4
4	*NAGPA*	c.139C>T	rs371054576	8056	Polymorphism		Neutral	Neutral	Neutral	B	Decreased stability	1
5	*NAGPA*	c.1394 C>T	rs7188856	0.3973	Polymorphism automatic		Neutral	Neutral	Low	PD	Decreased stability	1
Synonymous variants	**Synonymous variants**	**Synonymous variants**	**Synonymous variants**	**Synonymous variants**	**Synonymous variants**	**Synonymous variants**	**Synonymous variants**	**Synonymous variants**	**Synonymous variants**	**Synonymous variants**	**Synonymous variants**	**S v**
6	*GNPTAB*	**c.1932A>G**	rs10778148	0.4236	-	-	-	-	-	-	-	-
7	*GNPTG*	**c.702T>C**	rs532275192	23353	-	-	-	-	-	-	-	-
8	*GNPTG*	**c.813G>A**	rs377647926	64781	-	-	-	-	-	-	-	-
9	*NAGPA*	**c.333 G>A**	rs2972272	0.5259	-	-	-	-	-	-	-	-
10	*NAGPA*	**c. 1485C>T**	rs887854	0.7803	-	-	-	-	-	-	-	-
Non coding variants	**Non coding variants**	**Non coding variants**	**Non coding variants**	**Non coding variants**	**Non coding variants**	**Non coding variants**	**Non coding variants**	**Non coding variants**	**Non coding variants**	**Non coding variants**	**Non coding variants**	**N c i v a**
11	*GNPTG*	**-4 C>T**	rs554707396	69074	-	-	-	-	-	-	-	-
12	*NAGPA*	**c.1174+53C>A**	rs2237412	0.7424	-	-	-	-	-	-	-	-

DANN score 1: most damaging;

D: Damaging; T: Tolerant; PD: Possibly Damaging; B: Benign; VUS: Variant with Uncertain Significance

Table 3: Variant profile for the three putative genes for stuttering in probands with pathogenic mutation

	GNPTAB	Zygosity	GNPTG	Zygosity	NAGPA	Zygosity
STU 29	**c.3598G>A (E19)**	Het	-	-	**c.333 A>G (E2)**	Het
	c.1932A>G (E 13)	Homo	-	-	**c. 1485C>T (E10)**	Homo
					c.1174+53C>A (I7)	Het
STU 63	**c.1932A>G (E 13)**	Homo	**c.802A>C (E10)**	Het	**c.333 A>G (E2)**	Homo
					c. 1485C>T (E10)	Homo
					c.1174+53C>A (I 7)	Homo
STU 34	**c.1932A>G (E 13)**	Homo	-	-	**c.131G>C**	Het
					c.333 A>G (E2)	Homo
					c. 1485C>T (E10)	Homo
					c.1174+53C>A	Het

Likely pathogenic variant Synonymous variant Non coding variant

Table 4: Segregation pattern and genotype-phenotype correlation of likely pathogenic variants identified in the three putative genes for stuttering among the 64 probands screened

Code	Individual	Age	Sex	Phenotype	Gene	Genotype	Remarks
STU-29	**STU-29**	**STU-29**	**STU-29**	**STU-29**	**STU-29**	**STU-29**	
II-2	Father	45	M	**Affected**	*GNPTAB*	**c.3598G>A**/+	**Cosegregatio of the pathogenic allele suggests a dominant inheritance pattern Familial non consan-guineous**
II-4	Mother	38	F	Unaffected		+/+	
III-4	Brother	20	M	**Affected**		**c.3598G>A**/+	
III-6	Sister	17	F	Unaffected		+/+	
III-7	**Proband**	16	M	**Affected**		**c.3598G>A**/+	
III-8	Younger brother	15	M	Unaffected		+/+	
STU-63	**STU-63**	**STU-63**	**STU-63**	**STU-63**	**STU-63**	**STU-63**	
II-5	Father	50	M	Unaffected	*GNPTG*	+/+	*de novo* **variation Sporadic non consan-guineous**
II-10	Mother	45	F	Unaffected		+/+	
III-1	**Proband**	24	M	**Affected**		**c.802A>C**/+	
III-2	Sister	19	F	Unaffected		+/+	
STU-34	**STU-34**	**STU-34**	**STU-34**	**STU-34**	**STU-34**	**STU-34**	
II-5	Father	50	M	Unaffected	*NAGPA*	**c.131G>C**/+	**The variation does not cosegre-gate with affected status Sporadic non consan-guineous**
II-6	Mother	40	F	Unaffected		+/+	
III-3	Brother	18	M	Unaffected		**c.131G>C**/+	
III-4	**Proband**	14	M	**Affected**		**c.131G>C**/+	
III-5	Younger brother	11	M	Unaffected		DNA unavailable	

indicates likely pathogenic missense variation

Table 5: Lysosomal enzyme study in the plasma of a stuttering family

Family STU 63	**Genotype of *GNPTG* gene**	**LYSOSOMAL ENZYMES**	**LYSOSOMAL ENZYMES**	**LYSOSOMAL ENZYMES**
		Arylsulfatase A (Normal Range 30–268 nmol/hr/ mg protein)	**Hexosaminidase – A (Normal Range 90-385nmol/hr/mg protein)**	**B- γαλαςτοσι-δασε (Normal Range 470–2500nmol/hr/mg protein)**
STU 63-1 (proband)	**c.802A>C/+**	32.6	106.9	581.6
STU 63-2 (father) unaffected	+/+	31.9	108.1	489.7
STU 63-3 (mother) unaffected	+/+	38.2	113.1	631.9
STU 63-4 (sister) unaffected	+/+	33.6	116.1	506.3

Likely pathogenic variant

Figure legends

Figure 1: The twelve specific exons screened across the three genes implicated in stuttering

Figure 2a: Distribution of the variants identified in the three putative genes (*GNPTAB* , *GNPTG* and *NAGPA*) among the 64 probands with stuttering. Read clockwise starting with 12 O' clock position. ***Note:*** *Since some probands showed more than one variant the total sample size (64) will not tally.*

Figure 2b: Distribution of the variants identified in the three putative genes for stuttering

Figure 3: Partial chromatograms of the p.Glu1200Lys mutation (*GNPTAB*) segregating in a family with stuttering

Figure 4: Partial chromatogram of c.802A>C (p.Ile268Leu) variation in *GNPTG* gene

Figure 5: Partial chromatogram of c.131G>C (p.Arg44Pro) mutation in *NAGPA* gene

Figure 6: The relative levels of GNPTG, GNPTAB, and NAGPA mRNA expression were determined in WBC from blood sample of stuttering patients by real-time PCR normalized to β-actin expression. Data indicates [?]Ct values ±SD

List of Appendices

Figures

Figure A1: Partial chromatograms of *GNPTAB* variants observed in the study

Figure A2: Partial chromatograms of *GNPTG* variants observed in the study

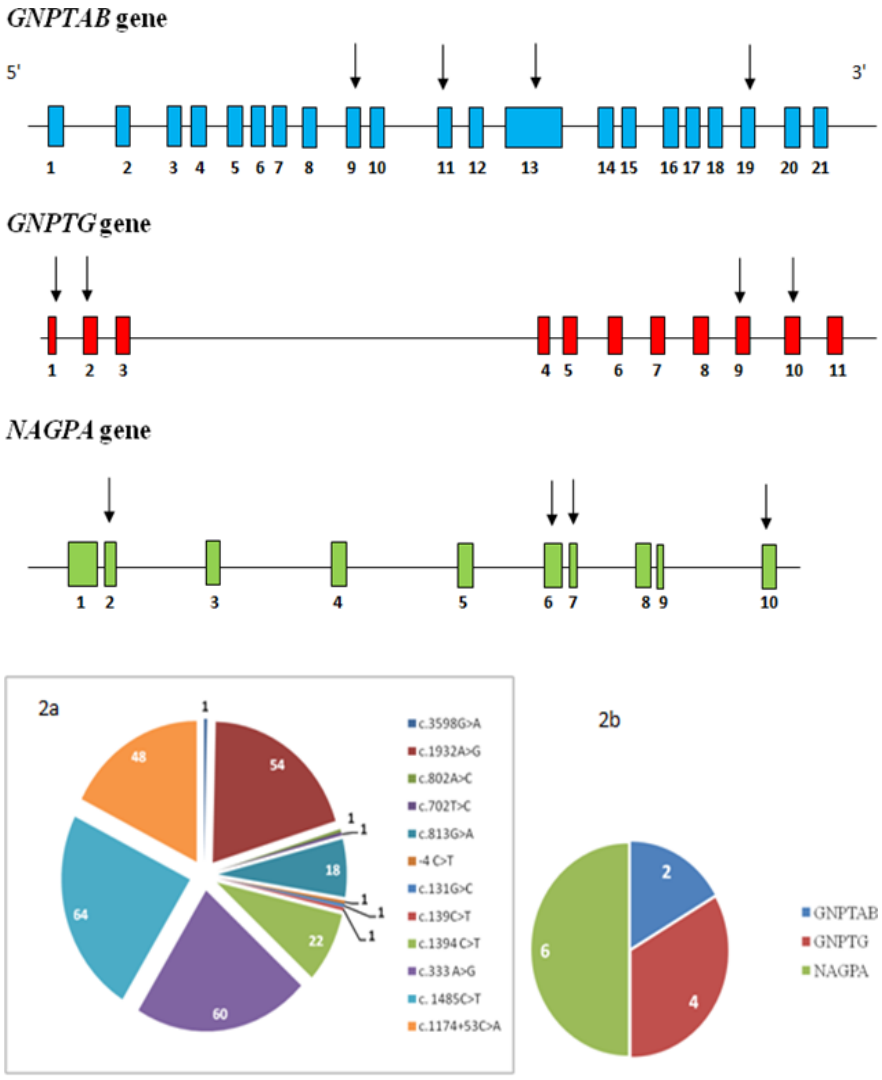
Figure A3: Partial chromatograms of variants observed in *NAGPA*gene

Figure A4: Alignment of native and mutated secondary structure of GNPTAB protein identified the loss of helix and addition of turn at the site of mutation.

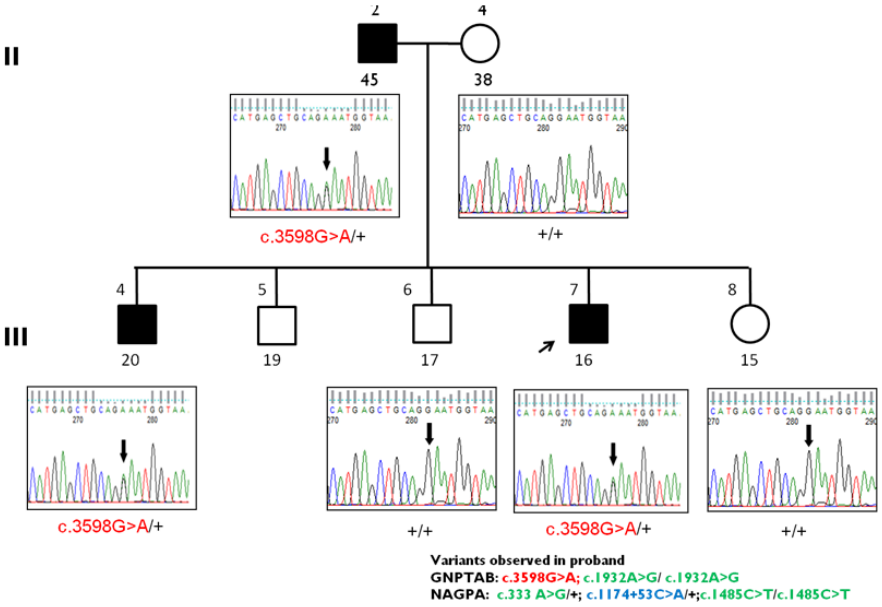
Tables

Table A1 : Demogenetic details of 64 probands with stuttering involved in mutation screening

Table A2: Real time nucleotide primer sequences of target (*GNPTAB* , *GNPTG* , *NAGPA*) and endogenous (β-*actin*) genes

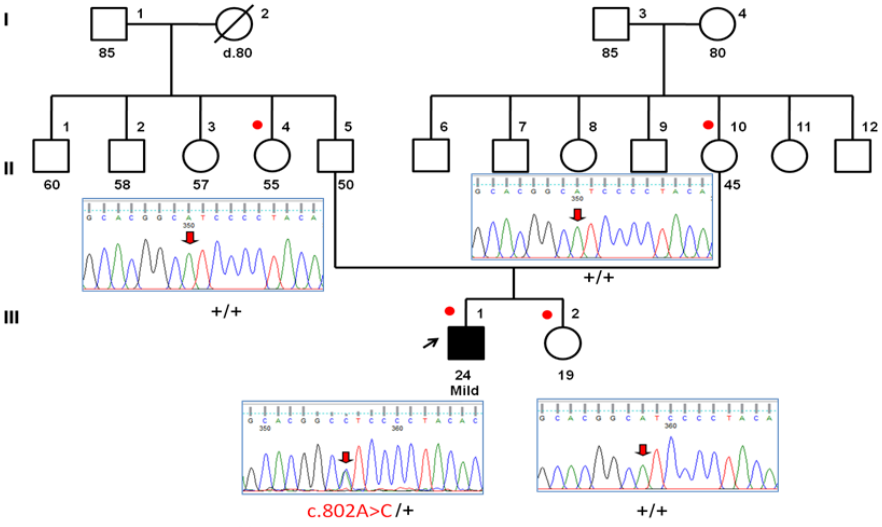


FAMILY STU 29



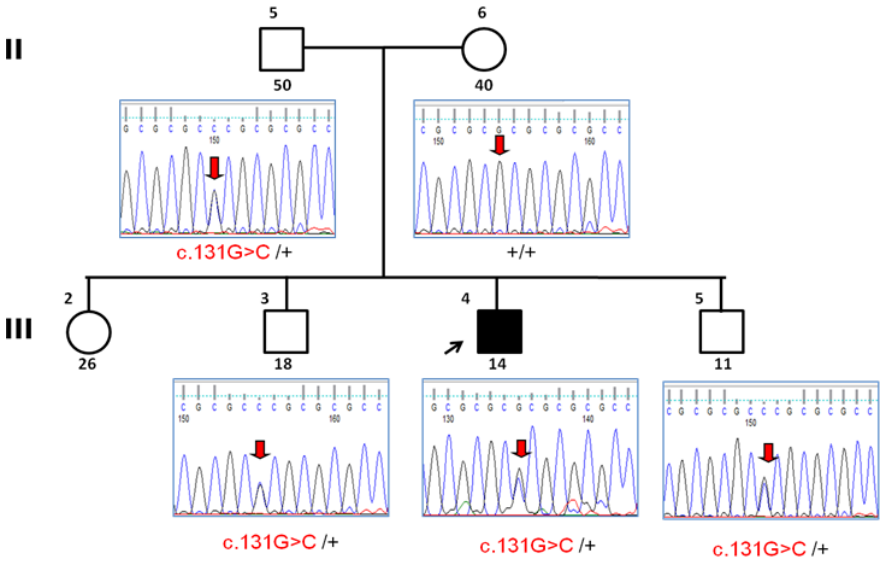
Variants observed in proband
GNPTAB: c.3598G>A; c.1932A>G/ c.1932A>G
NAGPA: c.333 A>G/+; c.1174+53C>A/+;c.1485C>T/c.1485C>T

FAMILY STU 63

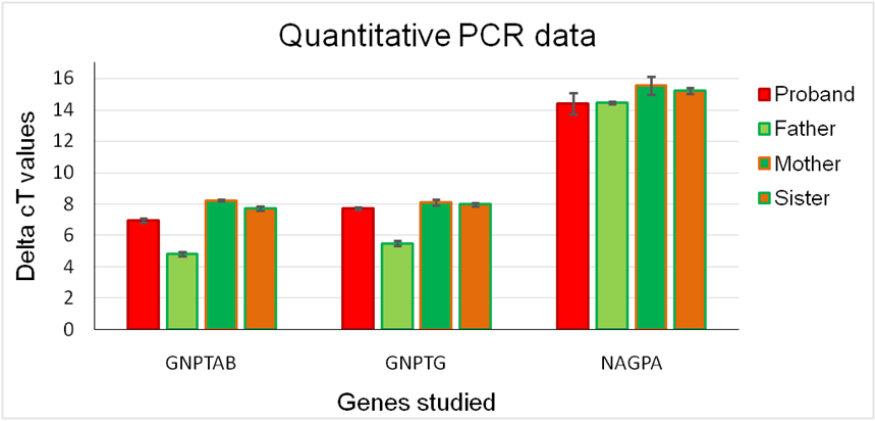


Variants observed in proband
GNPTAB: c.1932A>G/ c.1932A>G
GNPTG: c.802A>C/+
NAGPA: c.333 A>G/c.333 A>G; c.1174+53C>A/ c.1174+53C>A; c.1485C>T/c.1485C>T

FAMILY STU 34

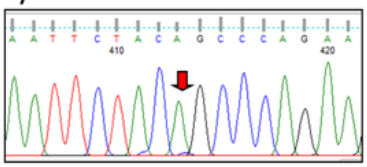


Variants observed in proband
GNPTAB: c.1932A>G/c.1932A>G
NAGPA: c.131G>C/+; c.333 A>G/c.333 A>G; c.1174+53C>A/+; c.1485C>T/c.1485C>T

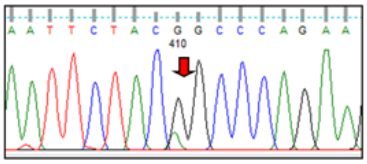


SYNONYMOUS
c.1932A>G; Thr644Thr
Exon 13 rs10778148

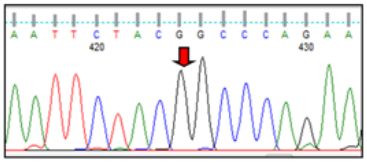
+/+



T644T/+

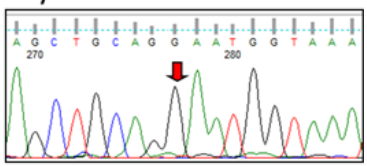


T644T/T644T

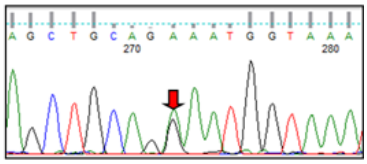


MISSENSE
c.3598G>A ; Glu1200Lys
Exon 19 rs137853825

+/+

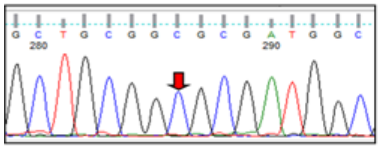


E1200K/+

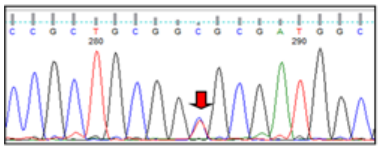


NON CODING
-4 C>T; 5'UTR
Exon 1 rs554707396

+/+

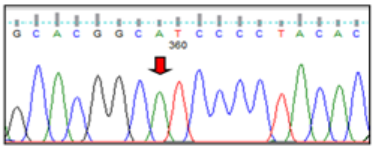


-4 C>T/+

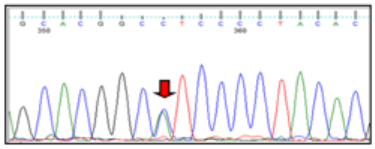


MISSENSE
c.802A>C; Ile268Leu
Exon 10 rs759796840

+/+

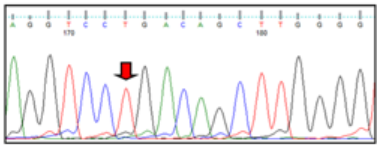


I268L/+

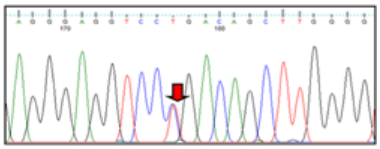


SYNONYMOUS
c.702T>C; Pro234Pro
Exon 9 rs532275192

+/+

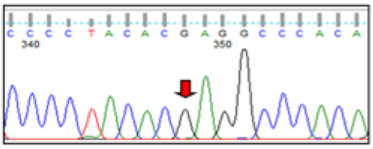


P234P/+



SYNONYMOUS
c.813G>A; Thr271Thr
Exon10 rs377647926

+/+



T271T /+