Genomic Profiling of Adult Onset Isolated Focal Dystonia in a Group of Romanian Patients

R. Cocoş[1] , I. Popescu-Olaru[2,3], O.Băjenaru [2,5], L. Cozma[2,3], L.Dumitescu[2,3], R. Tănăsescu[2,3,5], Florina Raicu[1,6] and B. O. Popescu [2,3,7]

[1] "Carol Davila" University of Medicine and Pharmacy – Department of Medical Genetics Bucharest, [2]"Carol Davila" University of Medicine and Pharmacy – Department of Neurosciences Bucharest, [3]Colentina Clinical Hospital – Department of Neurology Bucharest, [4]" Ana Aslan" National Institute of Geriatrics and Gerontology Bucharest, [5] University of Nottingham – Academic Clinical Neurology, Division of Clinical Neuroscience UK, [6] Francisc I. Rainer Anthropological Research Institute, Romanian Academy Bucharest, [7]"Victor Babes" National Institute of Pathology – Laboratory of Ultrastructural Pathology, Bucharest Romania

Introduction

Background

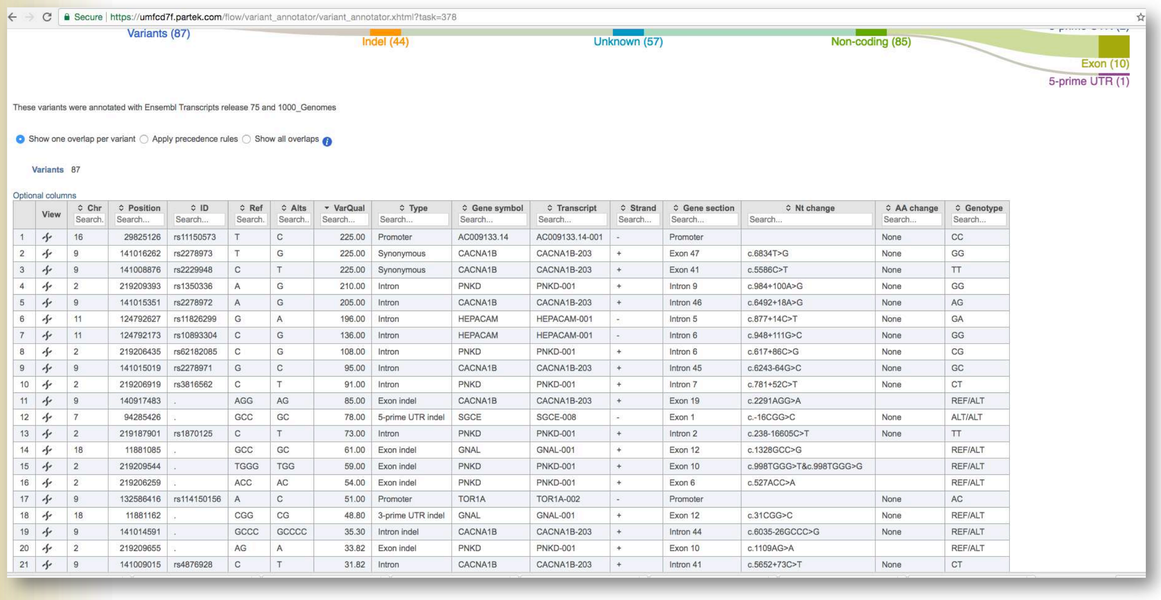
The recent evolution in the development of Next-Generation Sequencing (NGS) techniques and the growing number of candidate genes hopefully help to identify more dystonia-related mutations and risk variants allowing the improvement of the knowledge on the molecular basis of movement disorders. Genetic factors play an important role in isolated dystonia considering that more than 10 percentages of affected persons have a positive familial history. Targeted NGS has shown considerable potential in testing disease-targeted gene panels, as opposed to the whole exome and genome sequencing, and has proved a greater depth of coverage that increases the detection of indels that might be missed by exome sequencing and reduces the need to interpret sequence variants.

The presence of various phenotype-genotype correlations characterizing the isolated dystonia syndromes indicates the potential coexistence of several synergic genetic and/or environmental factors. The screening approach using the targeted next generation sequencing of a customized gene panel can represent an adequate method with clinical relevance beeing in the same time a promising method for identification of potential biologically relevant markers considering that we will include also the promoters, untranslated and intronic regions, which were mostly ignored by the previously published articles.

Aim

We proposed to identify the relevant genomic profiles of adult-onset isolated focal dystonia by gaining insights into the interplay of potential causative disease-associated genetic variants at the genomic level. Isolated focal dystonia is the third most common movement disorder after Parkinson's disease and essential tremor. The etiology is genetic in some cases, but over 75% are idiopathic.

Fig. 1 – Variants list



	View	Chr	Position	ID	Ref	Alts	VarQual	Type	Gene symbol	Transcript	Strand	Gene section	Nt change	AA change	Genotype
1		16	29825126	rs11150573	T	C	225.00	Promoter	AC009133.14	AC009133.14-001	-	Promoter		None	CC
2		9	141016262	rs2278973	T	G	225.00	Synonymous	CACNA1B	CACNA1B-203	+	Exon 47	c.6834T>G	None	GG
3		9	141008876	rs2229948	C	T	225.00	Synonymous	CACNA1B	CACNA1B-203	+	Exon 41	c.5586C>T	None	TT
4		2	219209393	rs1350336	A	G	210.00	Intron	PNKD	PNKD-001	+	Intron 9	c.984+100A>G	None	GG
5		9	141015351	rs2278972	A	G	205.00	Intron	CACNA1B	CACNA1B-203	+	Intron 46	c.6492+18A>G	None	AG
6		11	124792627	rs11826299	G	A	196.00	Intron	HEPACAM	HEPACAM-001	-	Intron 5	c.877+14C>T	None	GA
7		11	124792173	rs10893304	C	G	136.00	Intron	HEPACAM	HEPACAM-001	-	Intron 6	c.948+111G>C	None	GG
8		2	219206435	rs62182085	C	G	108.00	Intron	PNKD	PNKD-001	+	Intron 6	c.617+86C>G	None	CG
9		9	141015019	rs2278971	G	C	95.00	Intron	CACNA1B	CACNA1B-203	+	Intron 45	c.6243-64G>C	None	GC
10		2	219206919	rs3816562	C	T	91.00	Intron	PNKD	PNKD-001	+	Intron 7	c.781+52C>T	None	CT
11		9	140917483	.	AGG	AG	85.00	Exon indel	CACNA1B	CACNA1B-203	+	Exon 19	c.2291AGG>A		REF/ALT
12		7	94285426	.	GCC	GC	78.00	5-prime UTR indel	SGCE	SGCE-008	-	Exon 1	c.-16CGG>C	None	ALT/ALT
13		2	219187901	rs1870125	C	T	73.00	Intron	PNKD	PNKD-001	+	Intron 2	c.238-16605C>T	None	TT
14		18	11881085	.	GCC	GC	61.00	Exon indel	GNAL	GNAL-001	+	Exon 12	c.1328GCC>G		REF/ALT
15		2	219209544	.	TGGG	TGG	59.00	Exon indel	PNKD	PNKD-001	+	Exon 10	c.998TGGG>T&c.998TGGG>G		REF/ALT
16		2	219206259	.	ACC	AC	54.00	Exon indel	PNKD	PNKD-001	+	Exon 6	c.527ACC>A		REF/ALT
17		9	132586416	rs114150156	A	C	51.00	Promoter	TOR1A	TOR1A-002	-	Promoter		None	AC
18		18	11881162	.	CGG	CG	48.80	3-prime UTR indel	GNAL	GNAL-001	+	Exon 12	c.31CGG>C	None	REF/ALT
19		9	141014591	.	GCCC	GCCCC	35.30	Intron indel	CACNA1B	CACNA1B-203	+	Intron 44	c.6035-26GCCC>G	None	REF/ALT
20		2	219209655	.	AG	A	33.82	Exon indel	PNKD	PNKD-001	+	Exon 10	c.1109AG>A		REF/ALT
21		9	141009015	rs4876928	C	T	31.82	Intron	CACNA1B	CACNA1B-203	+	Intron 41	c.5652+73C>T	None	CT

30-Genes Panel Design

This panel of genes was online designed by us in three separate PCR primer pools covering the coding exons and 50bp of introns, promoters and the UTRs of the 30 genes using Ion *AmpliSeq Designer v3.0* according the Ion Torrent recommendations. The raw data from the sequencing runs were analysed using *Torrent Suite* analysis pipeline and a customized pipeline. The panel includes known and newly identified candidate genes involved in isolated dystonia. For the present study, the 30 nuclear genes were initially analysed with the possibility of expanding or redesigning primer panel to improved the uniformity of coverage.

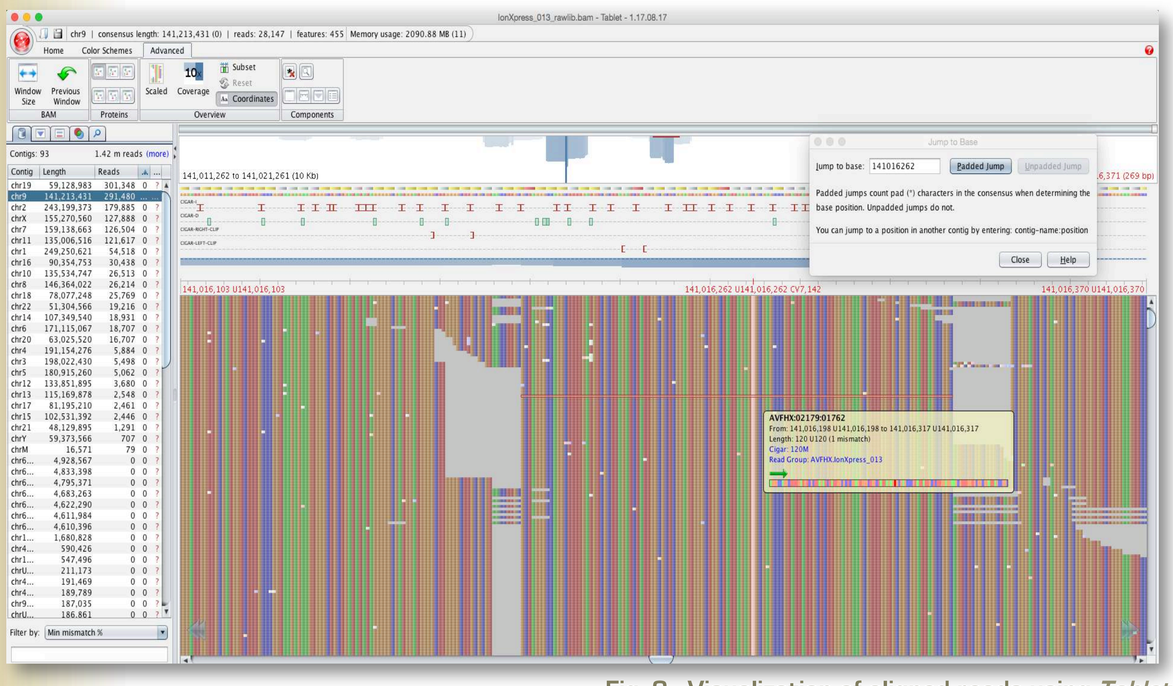


Fig. 2 - Visualization of aligned reads using *Tablet*

Conclusion

The subsequent analysis of negative-known DYTs patients for the genes in the customized panel using the Whole Exome Sequencing could offer the opportunity of uncovering de novo variants causing-disease. Analyzing of a large case group might point out to the reducing of the number of genes in the AmpliSeq panel, thus narrowing down the number of genes to those with a real clinical effect. Our approach not only will provide insights into identifying new causative mutations and SNPs, but also could prove to be a promising method to find biologically relevant new key molecules that might be pharmaceutically targeted. A special attention will be given on some specific synonymous mutations founded with high frequency in a representative number of patients which could influence protein folding and function. This approach will give the geneticists the ability to increase clinical sensitivity for many existing tests and will have a substantial contribution for identifying of unique and rare variations associated to dystonia, which can be assayed at a large scale. The use of an appropriate multigene panel could become a valuable tool for dystonia diagnosis and for appropriate management and treatment.

Research Funding

The authors declare no conflict of interest.

This work was supported by the CNCS-UEFISCDI grant PN-III-PID-PCE-2016-0696.

Materials and Methods

DNA extraction

DNA was extracted from buccal cells with PureLink Genomic DNA kit according to the protocol of the manufacturer. Swab samples from 120 affected patients that meet the inclusion criteria. Inclusion criteria were applied following the diagnosis of adult-onset isolated focal dystonia according to the European/ EFNS guidelines on the diagnosis and treatment of primary dystonia.

Patients selection

In the first stage of our study, we analyzed onto the Ion Torrent PGM platform a well-described homogenous group of 120 Romanian dystonia cases using our own designed Targeted Sequencing AmpliSeq panel of 30 genes previously found to be associated with dystonia (DYT) and other movement disorders. All samples had 97.85% average of target regions with coverage by at least 30 folds. Data were analyzed using a bioinformatics in house pipeline based on SAMtools, GATK, Picard and Annovar software for reading alignments, variant calling, filtering and annotation. We run the same data using Partek Flow software for concordance in variants detection.

Fig. 3 – Run Report Summary

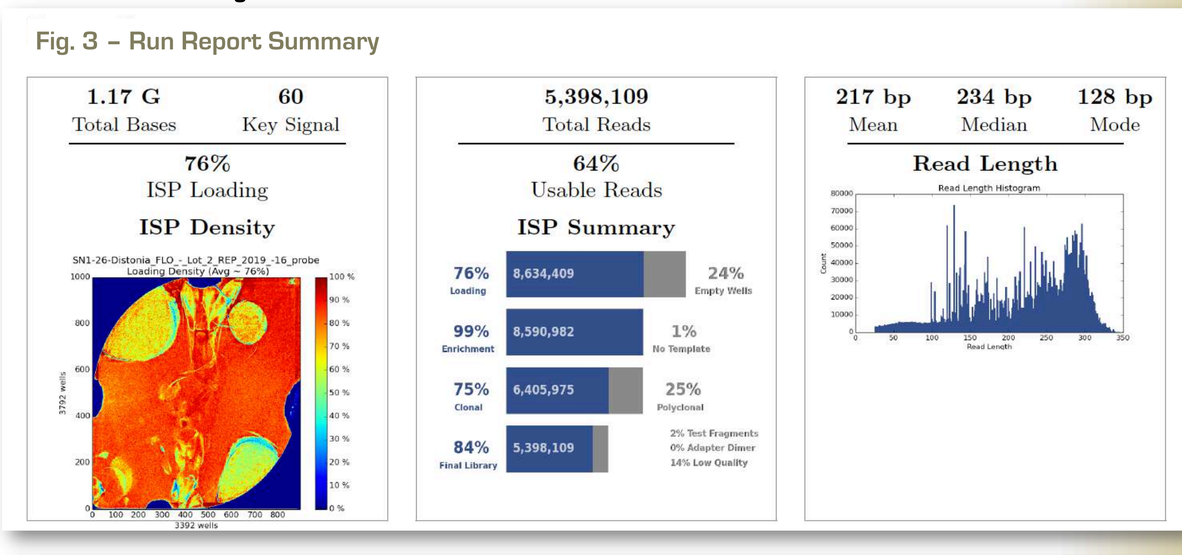
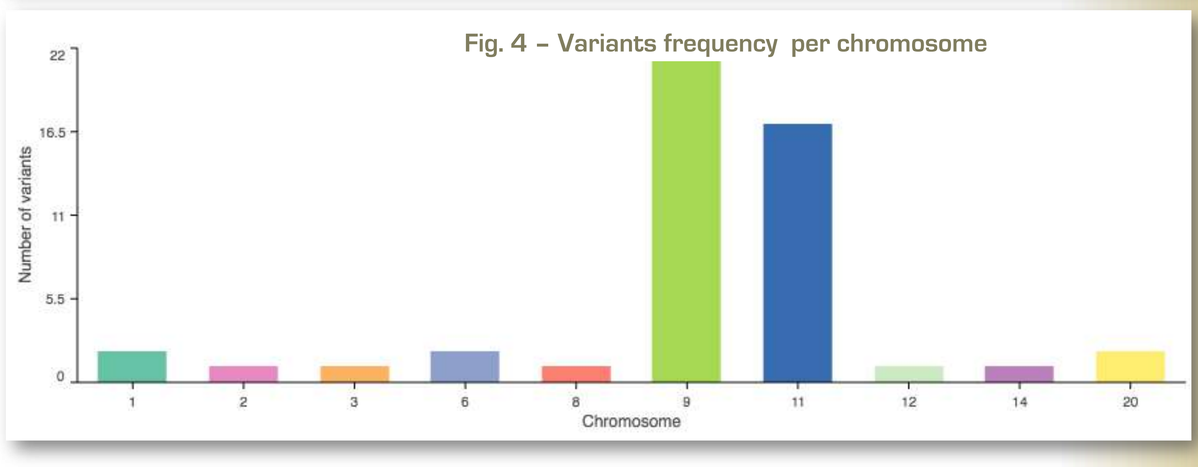


Fig. 4 – Variants frequency per chromosome



Next Generation Sequencing

For the library preparation *Ion AmpliSeq Library Kit v2.0* was used and subsequently to the ligation of adaptors with *Ion Xpress Barcode Adapters Kit* we cleaned the libraries *using Agencourt AMPure XP reagent*. The size distribution of the DNA fragments was analysed on the Agilent Bioanalyzer using the High Sensitivity Kit.

After the preparation of libraries to equimolar amounts, the library pool was subjected to emulsion PCR using *Ion PGM Template OT 2 200 View Kit*. We evaluated the efficiency of the emulsion PCR by estimation of the ISPs (Ion Sphere Particles) with DNA using the Qubit Ion Sphere Quality Control Kit. Enriched ISPs on Ion OneTouch ES was loaded onto an Ion 318™ Chip Kit v2 and sequence on the Ion Torrent Personal Genome Machine (PGM) with the 200-bp single-end run configuration using *Ion PGM HI-Q View Sequencing Kit* was performed.

In addition, for a systematic comparison we assessed the impact of different variant calling pipelines comparing the same read aligner, BWA-MEM, and four variant callers GATK-HaplotypeCaller, SAMtools (the vcfutils.pl script), Platypus and FreeBayes.

Confirmatory Sanger sequencing will be performed using direct Sanger sequencing on ABI PRISM 3130XL Sequencer using the Big Dye Terminator v3.1 Cycle Sequencing kit. Analyses of the sequencing data will be performed using the Seqscape software Version 2.5.

Results

Molecular genetic screening identified mutations in 96/120 of the clinically diagnosed adult-onset isolated focal dystonia subjects. All identified mutations were previously reported to be associated with DYTs.

All samples had 98.37% average of target regions with coverage by at least 30 folds. The mean read length was approximately 210-270 bp. Data were analyzed also using a bioinformatics in house pipelines software for reading alignments, variant calling, filtering and annotation.

We identified 87 possible causative variants in 10 different genes, mainly located on chromosomes 9 and 11, excluding common polymorphisms, and a higher concordance among the GATK, SAMtools and Platypus variant calling pipelines compared to a lower concordance with the FreeBayes variant caller.

Our data found an essential variation between the analyzed variant-calling pipelines and suggests the standardization of the available bioinformatics pipelines to reduce the implication on analyzing the next-generation sequencing data.

ESHG
EUROPEAN SOCIETY OF HUMAN GENETICS
European Human Genetics Conference
Gothenburg, Sweden | June 15-18, 2019