This study was supported by FCT project HMSP-ICT/0016/2013

SVInterpreter: a web-based tool for structural variants inspection and identification of possible disease-causing candidate genes

J. Fino[1], B. Marques[1], Z. Dong[2,3,4], D. David[1]

https://dgrctools-insa.min-saude.pt/cgi-bin/SVInterpreter.py

[1]Department of Human Genetics, National Health Institute Doutor Ricardo Jorge, Lisbon, Portugal
[2]Department of Obstetrics and Gynecology, The Chinese University of Hong Kong, Hong Kong, China
[3]Shenzhen Research Institute, The Chinese University of Hong Kong, Shenzhen, China
[4]Hong Kong Hub of Pediatric Excellence, The Chinese University of Hong Kong, Hong Kong, China

Introduction

- Current genomic sequencing technologies revolutionized the identification of balanced, unbalanced and complex structural variants (SV).

- Currently, due to the large number of identified variants per individual and the substantial revision of dispersed data that this entails, ascertainment of the pathogenicity and mechanistic links between SVs and human pathologies is a daunting task .

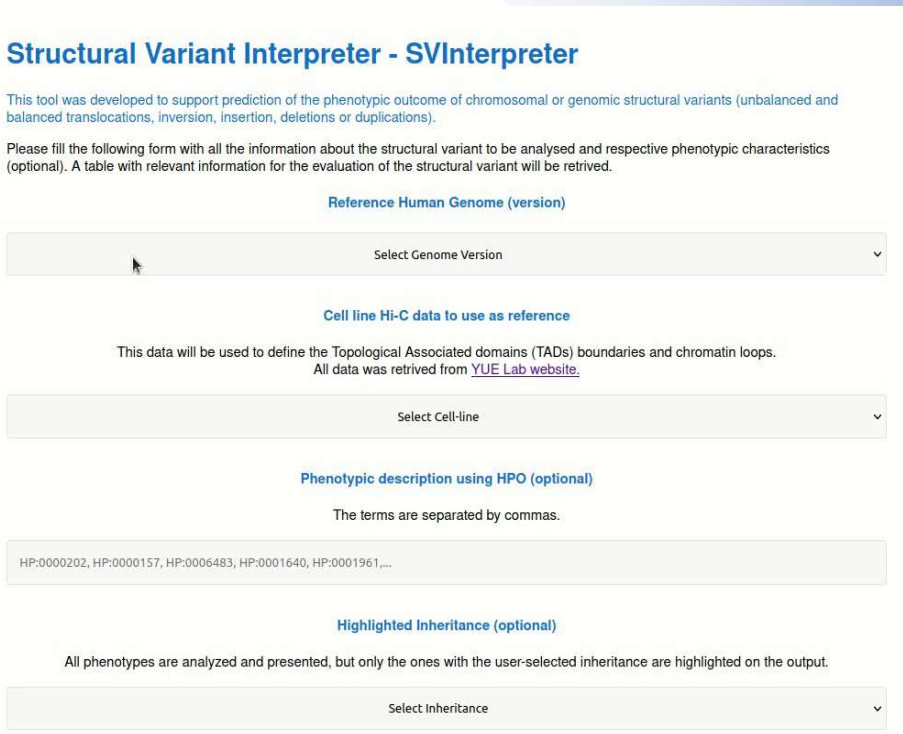
- Available tools are unable to gather the full information needed for SV evaluation.

To address the need of a comprehensive application to assist evaluation of clinical outcome of SVs, we present SVInterpreter.

https://dgrctools-insa.min-saude.pt/cgi-bin/SVInterpreter.py

Methods

- Python-CGI developed Web application, that uses local data and several API connections, including Ensembl and MARRVEL.
- Analyze SVs using Topologically Associated Domains as genome units.
- Compiles genomic data, including medically actionable genes, virtual gene panels, gene-associated disorders and respective phenotypic overlap, GeneHancer cluster of interactions, chromatin loops, among other.
- For CNVs, database overlap search and ACMG criteria are also calculated and presented.
- Retrieves an xlsx table with the compiled data, phenotypic overlap results, and the breakpoint location using ISCN 2020 nomenclature.



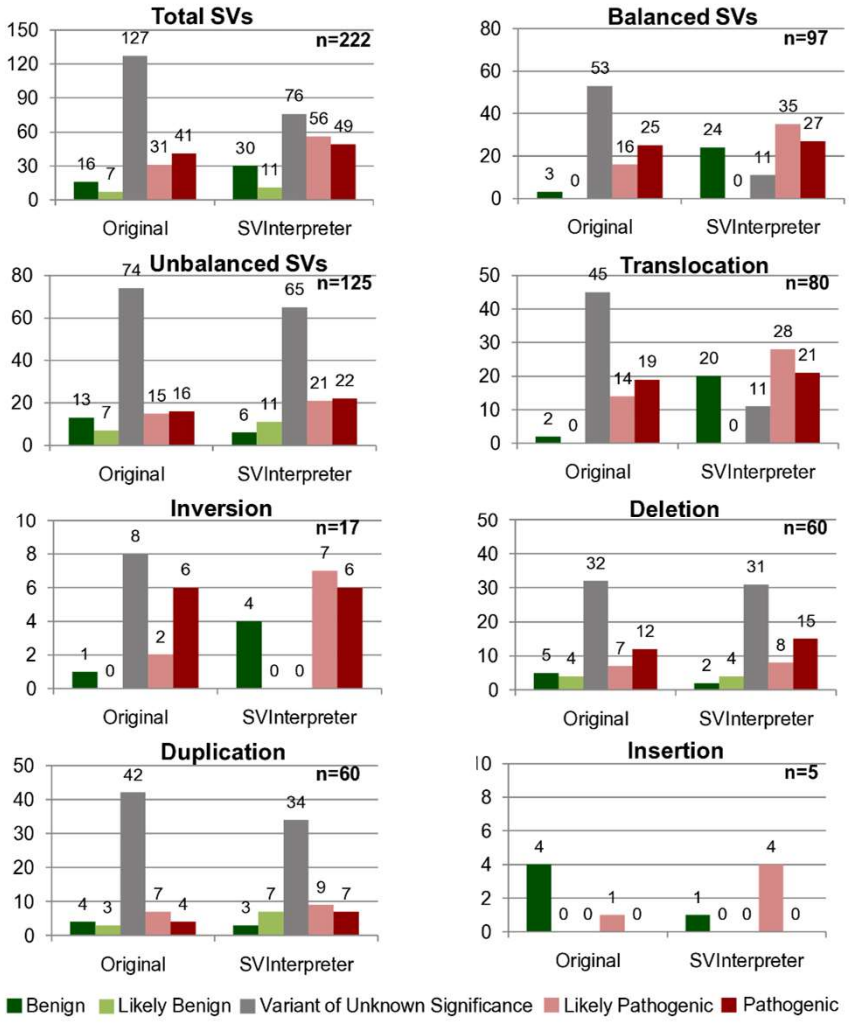
SVInterpreter input form layout.

https://dgrctools-insa.min-saude.pt/cgi-bin/SVInterpreter.py

Results

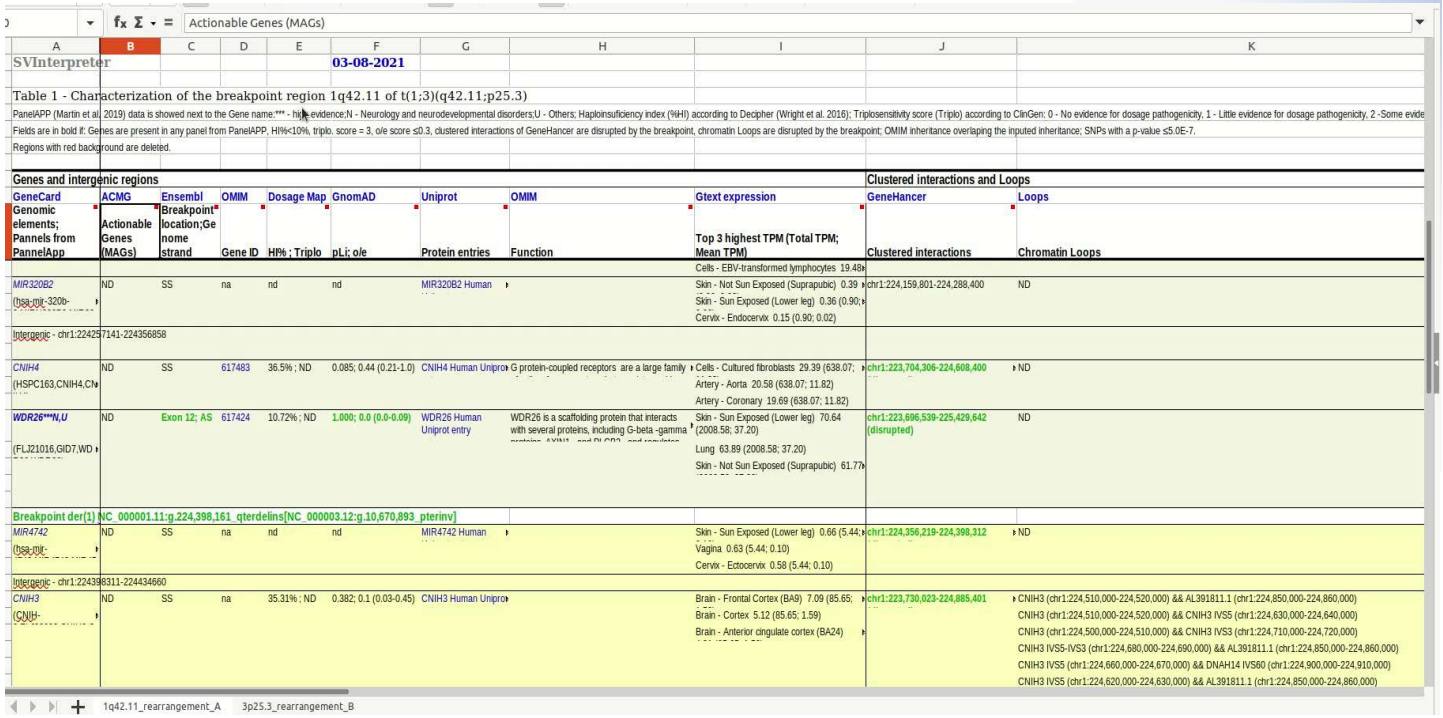
- We reevaluated 222 published SVs and 20 clinical cases analyzed by chromosomal microarray or genome sequencing.
- Our results corroborated more than half of the original predictions, decreased by 40% the variants of unknown significance, and indicated several potential position effect events.
- Using SVInterpreter data, as gene function, and animal models, we were able to indicate potential candidate genes not identified by any other approach
- For all clinical cases, original clinical outcome prediction was quickly confirmed, without any other data resource.

https://dgrctools-insa.min-saude.pt/cgi-bin/SVInterpreter.py



Conclusion

- To our knowledge, SVInterpreter is the most comprehensive TAD based tool to assist SV clinical outcome prediction.
- SVInterpreter resources as phenotypic overlap, indication of position effects (loop and GeneHancer interaction disruption), CNV overlap and classification criteria, greatly facilitated the SV evaluation.
- Based on gathered information, identification of possible disease-causing candidate genes and SVs is easily achievable.



Actionable Genes (MAGs)

SVInterpreter 03-08-2021

Table 1 - Characterization of the breakpoint region 1q42.11 of t(1;3)(q42.11;p25.3)

PanelAPP (Martin et al, 2019) data is showed next to the Gene name:*** - high evidence;N - Neurology and neurodevelopmental disorders;U - Others; Haploinsuficiency index (%HI) according to Decipher (Wright et al. 2016); Triplosensitivity score (Triplo) according to ClinGen: 0 - No evidence for dosage pathogenicity, 1 - Little evidence for dosage pathogenicity, 2 -Some evide

Fields are in bold if: Genes are present in any panel from PanelAPP, HI%<10%, triplo. score = 3, o/e score ≤0.3, clustered interactions of GeneHancer are disrupted by the breakpoint, chromatin Loops are disrupted by the breakpoint; OMIM inheritance overlaping the inputed inheritance; SNPs with a p-value ≤5.0E-7.

Regions with red background are deleted.

Genes and intergenic regions									Clustered interactions and Loops	
GeneCard Genomic elements; Pannels from PannelApp	ACMG Actionable Genes (MAGs)	Ensembl Breakpoint location;Genome strand	OMIM Gene ID	Dosage Map HI% ; Triplo	GnomAD pLi; o/e	Uniprot Protein entries	OMIM Function	Gtext expression Top 3 highest TPM (Total TPM; Mean TPM)	GeneHancer Clustered interactions	Loops Chromatin Loops
								Cells - EBV-transformed lymphocytes 19.48		
MIR320B2 (hsa-miR-320b-	ND	SS	na	nd	nd	MIR320B2 Human		Skin - Not Sun Exposed (Suprapubic) 0.39 Skin - Sun Exposed (Lower leg) 0.36 (0.90; Cervix - Endocervix 0.15 (0.90; 0.02)	chr1:224,159,801-224,288,400	ND
Intergenic - chr1:224257141-224356858										
CNIH4 (HSPC163,CNIH4,CI	ND	SS	617483	36.5% ; ND	0.085; 0.44 (0.21-1.0)	CNIH4 Human Uniprot	G protein-coupled receptors are a large family	Cells - Cultured fibroblasts 29.39 (638.07; Artery - Aorta 20.58 (638.07; 11.82) Artery - Coronary 19.69 (638.07; 11.82)	chr1:223,704,306-224,608,400	ND
WDR26***N,U (FLJ21016,GID7,WD	ND	Exon 12; AS	617424	10.72% ; ND	1.000; 0.0 (0.0-0.09)	WDR26 Human Uniprot entry	WDR26 is a scaffolding protein that interacts with several proteins, including G-beta -gamma	Skin - Sun Exposed (Lower leg) 70.64 (2008.58; 37.20) Lung 63.89 (2008.58; 37.20) Skin - Not Sun Exposed (Suprapubic) 61.77	chr1:223,696,539-225,429,642 (disrupted)	ND
Breakpoint der(1) NC_000001.11:g.224,398,161_qterdelins[NC_000003.12:g.10,670,893_pterinv]										
MIR4742 (hsa-miR-	ND	SS	na	nd	nd	MIR4742 Human		Skin - Sun Exposed (Lower leg) 0.66 (5.44; Vagina 0.63 (5.44; 0.10) Cervix - Ectocervix 0.58 (5.44; 0.10)	chr1:224,356,219-224,398,312	ND
Intergenic - chr1:224398311-224434660										
CNIH3 (CNIH-	ND	SS	na	35.31% ; ND	0.382; 0.1 (0.03-0.45)	CNIH3 Human Uniprot		Brain - Frontal Cortex (BA9) 7.09 (85.65; Brain - Cortex 5.12 (85.65; 1.59) Brain - Anterior cingulate cortex (BA24)	chr1:223,730,023-224,885,401	CNIH3 (chr1:224,510,000-224,520,000) && AL391811.1 (chr1:224,850,000-224,860,000) CNIH3 (chr1:224,510,000-224,520,000) && CNIH3 IVS5 (chr1:224,630,000-224,640,000) CNIH3 (chr1:224,500,000-224,510,000) && CNIH3 IVS3 (chr1:224,710,000-224,720,000) CNIH3 IVS5-IVS3 (chr1:224,680,000-224,690,000) && AL391811.1 (chr1:224,850,000-224,860,000) CNIH3 IVS5 (chr1:224,660,000-224,670,000) && DNAH14 IVS60 (chr1:224,900,000-224,910,000) CNIH3 IVS5 (chr1:224,620,000-224,630,000) && AL391811.1 (chr1:224,850,000-224,860,000)

1q42.11_rearrangement_A 3p25.3_rearrangement_B

SVInterpreter example output xlsx file of a t(1;3)

https://dgrctools-insa.min-saude.pt/cgi-bin/SVInterpreter.py