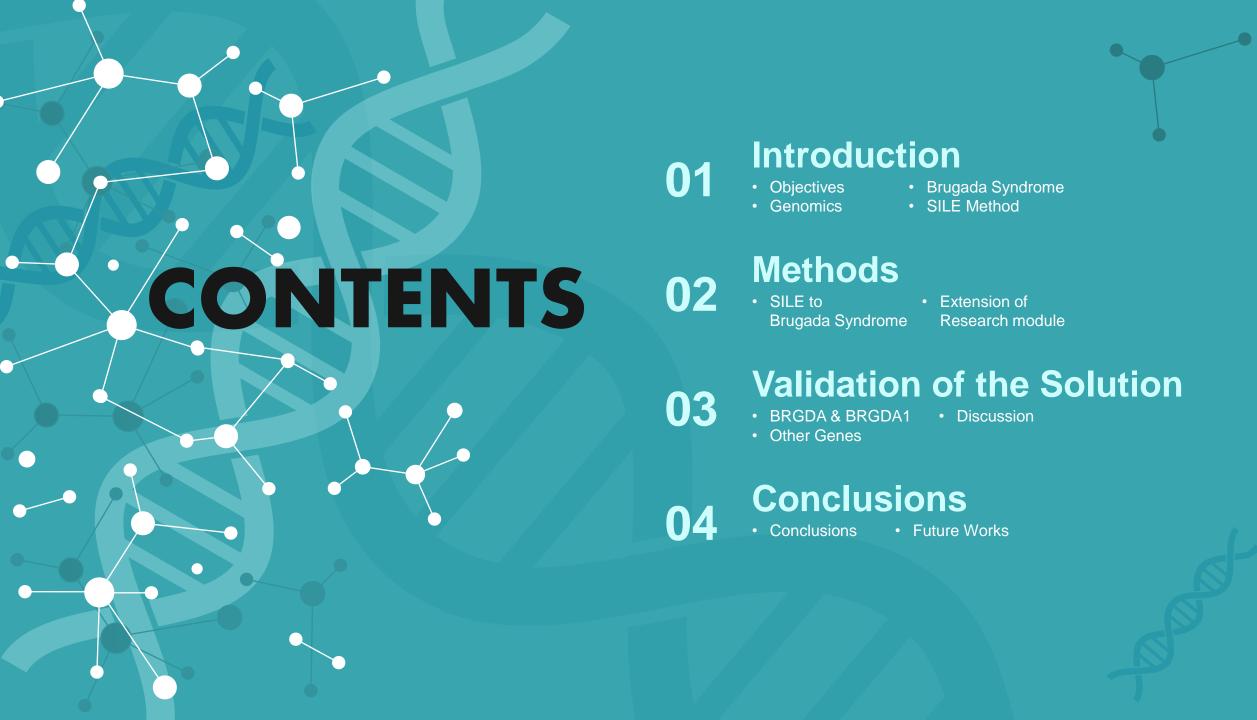
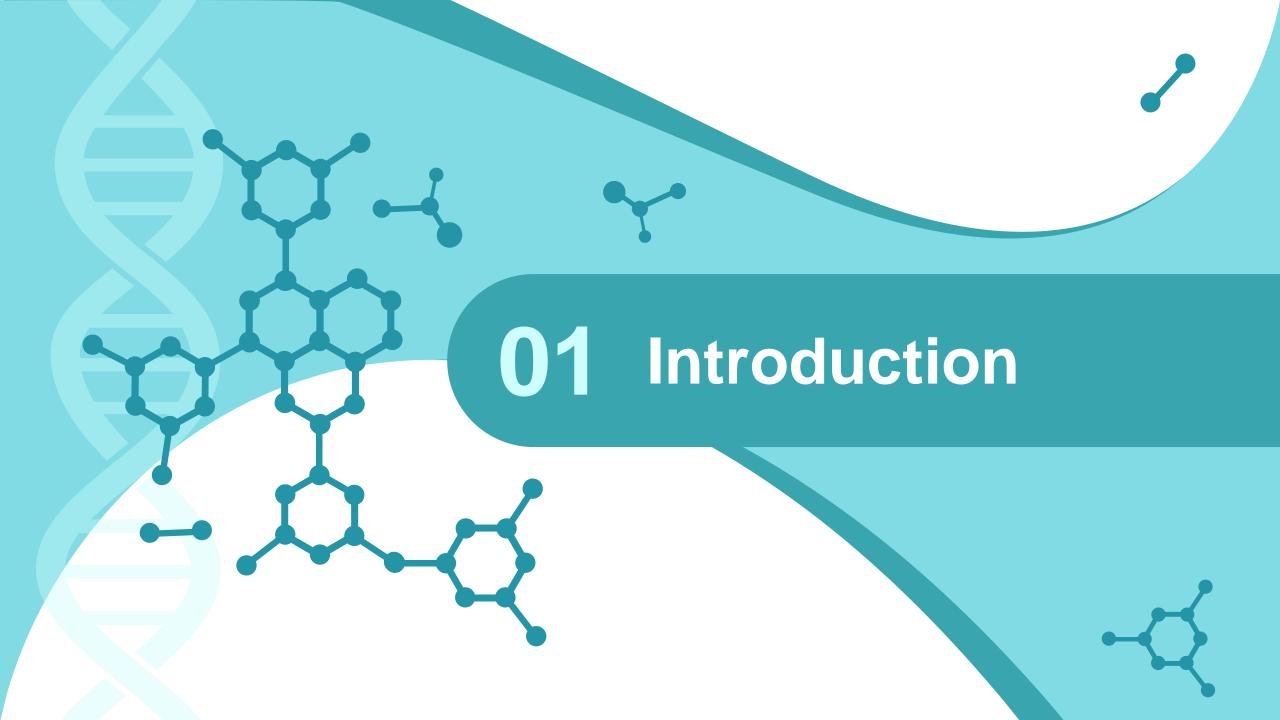




Identification of genomic variants using the SILE Method

Extension of the Research module & Focus on the Brugada Syndrome







INTRODUCTION

I) Objectives

- 1) Evaluation of SILE Method applied to Brugada Syndrome
- 2) Improvement of the Research module of SILE
- 3) Evaluation of the Improvement

II) SILE Method

S: Search L: Load

III) Brugada Syndrome

Diagnosis Treatment

IV) Genomics Concepts

DNA Variations
Chromosomes Genotype
Genes Phenotype

Genomics data are **scattered** in many data sources

Each of them containing **different** information about different genes and variants

Goal: collect as many relevant data as possible

There is **no** protocol / **method** for searching and identifing relevant information

SILE method to **integrate** all the relevant information in the same place



Research and **selection** of genomic data sets to have quality data

- **Filter** information gathered
- Discart non-relevant and duplicates
- Keep only variations that have sufficiently relevant evidence
- Manually solve inconsistencies
- Platform based on the CSHG (Conceptual Schema of the Human Genome)
- Validation carried out by the platform
- Extract knowledge from the information stored
- Platform to visualize data
- Comparison between genetic variations





Global Variome shared LOVD

SCN5A (sodium channel, voltage-gated, t

Curator: Global Variome, with Curator vacancy

Transcripts X Variants X Individuals X Diseases

Screenings >

Subm

The SCN5A gene homepage

General information			
Gene symbol	SCN5A		
Gene name	sodium channel, voltage-gated, type V, alpha subunit		
Chromosome	3		
Chromosomal band	p21		
Imprinted	Not imprinted		
Genomic reference	LRG 289		
Transcript reference	NM 198056.2		
Exon/intron information	NM 198056.2 exon/intron table		
Associated with diseases	ATFB10, BRGDA1, CMD1E, LQT3, PFHB1A, SIDS, SSS1,		
	<u>VF1</u>		
Citation reference(s)	-		
Refseq URL	Genomic reference sequence		
Curators (1)	Global Variome, with Curator vacancy		
Total number of public variants reported	<u>1639</u>		
Unique public DNA variants reported	<u>981</u>		
Individuals with public variants	<u>1509</u>		
Hidden variants	52		
Download all this gene's data	Download all data		
Notes	Establishment of this gene variant database (LSDB) was supported by the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement No 200754 - the GEN2PHEN project.		
Date created	April 29, 2010		
Date last updated	November 30, 2021		
Version	SCN5A:211130		

LOVD (Leiden Open Variation Database)

Open source database containing information about genomic variations

More than 23,000 different genes

Follows the reccomendations of **HGVS** (Human Genome Variation Society)

6 categories:

Genes	
Transcripts	
Variants	
Individuals	
Diseases	
Screenings	

BRUGADA

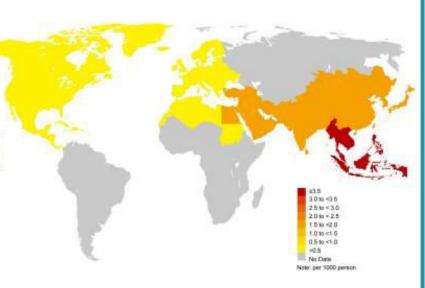
DIAGNOSIS

Rare cardiac arrhythmia

If left untreated can cause fainting, difficult breathing and high risk of **sudden cardiac death**, usually when sleeping

Can be caused by **mutation** in one gene:

Types	Associated Genes
Brugada syndrome 1 (BRGDA1)	SCN5A
Brugada syndrome 2 (BRGDA2)	GPD1L
Brugada syndrome 3 (BRGDA3)	CACNA1C
Brugada syndrome 4 (BRGDA4)	CACNB2
Brugada syndrome 5 (BRGDA5)	SCN1B
Brugada syndrome 6 (BRGDA6)	KCNE3
Brugada syndrome 7 (BRGDA7)	SCN3B
Brugada syndrome 8 (BRGDA8)	HCN4
Brugada syndrome 9 (BRGDA9)	KCND3



Globally affect 0,5 per 1000

In **Southeast Asia** 3,7 per 1000

TREATMENT

Symptomatic:

Implantation of Implantambel Cardioverterer-defibrillator (only treatment effective)

Asymptomatic:

Individual risk assesment Consider other risk factors (age, sex, external)

GENOMICS CONCEPTS

HUMAN GENOME:

DNA and RNA

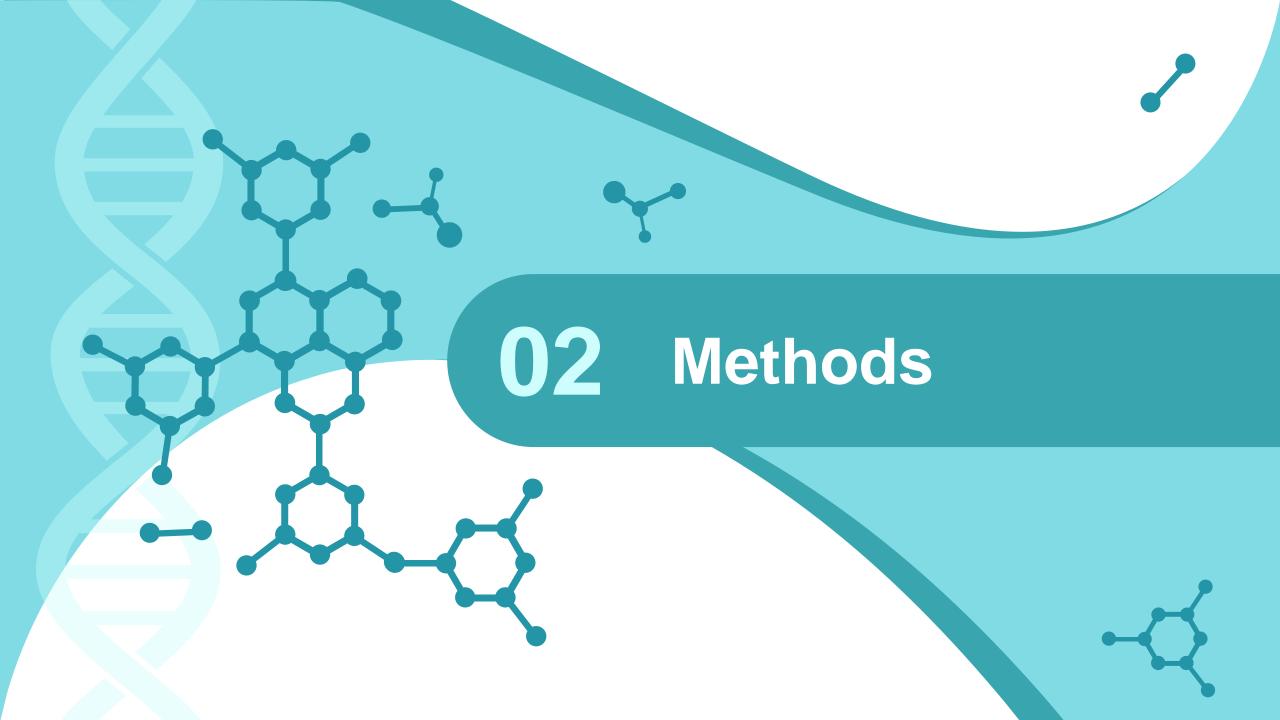
CHROMOSOMES

PHENOTYPE

GENES

GENOTYPE

VARIATIONS



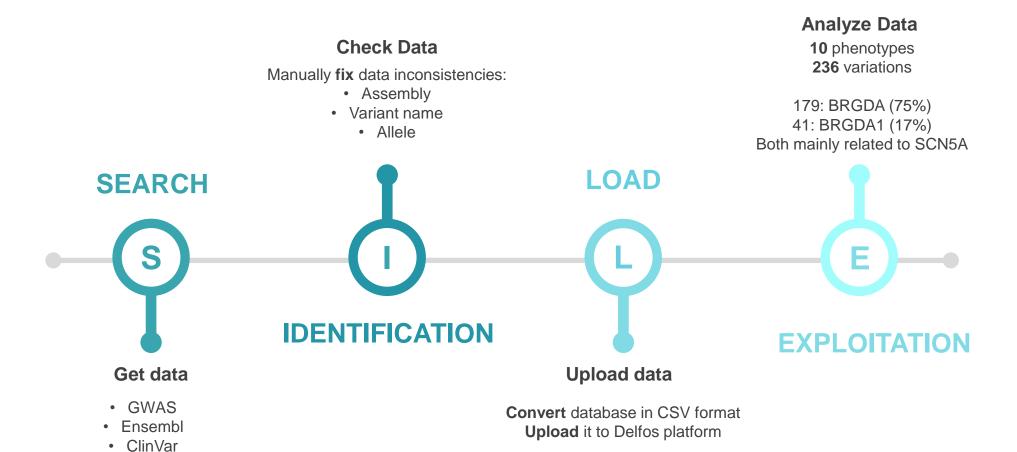


METHODS

1) Application of SILE method to BRUGADA Syndrome

2) Extension of Research module

SILE TO BRUGADA



SILE TO BRUGADA

Brugada Syndrome				
Chromosome Affected	Chr3: 38.649.686 - 38.548.060			
Gene affected	SCN5A			
Number of variations for the phenotype	179			
Evidence of variations	171 limited, 8 moderate			
Most common type of variation	Single nucleotide variant			

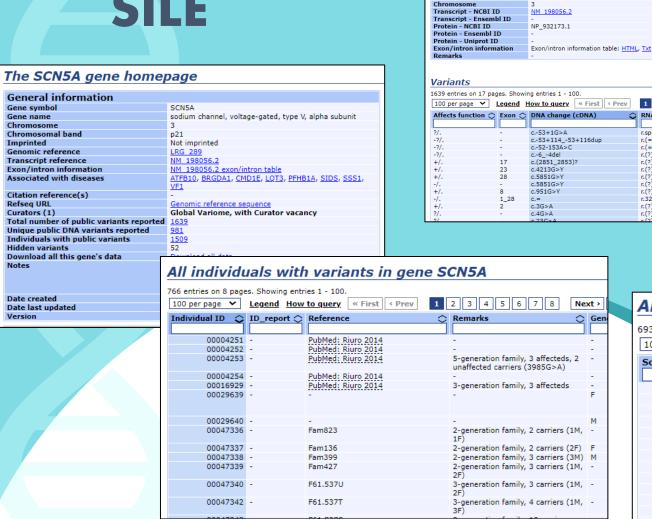
Brugada Syndrome 1				
Chromosome Affected	Chr3: 38.649.686 - 38.548.060			
Gene affected	SCN5A			
Number of variations for the phenotype	41			
Evidence of variations	7 limited, 34 moderate			
Most common type of variation	Single nucleotide variant			

1 variation for BRGDA2, BRGDA7, BRGDA8, BRGDA9

2 variations for BRGDA3, BRGDA5

7 variations for "Brugada Syndrome (Shorter-thannormal QT interval)"

1 variation for "Spontaneous Brugada pattern ECG"



Transcript #00018523 (NM_198056.2, SCN5A gene)

SCN5A (sodium channel, voltage-gated, type V, alpha subunit)

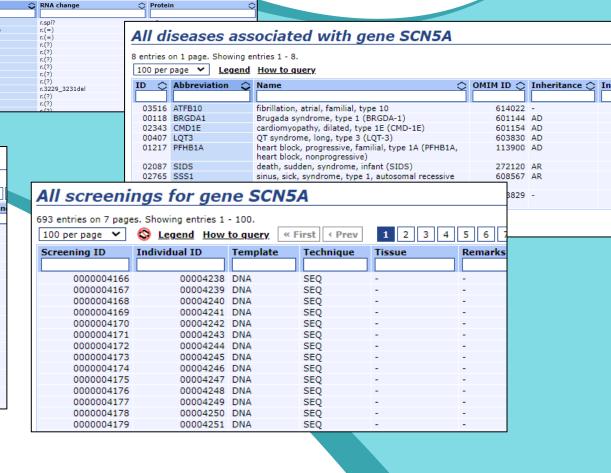
1 2 3 4 5 6 7 8 9 10 11 ...

Transcript name

Gene name

Rselenium library used for:

Transcripts, Individuals, Screenings

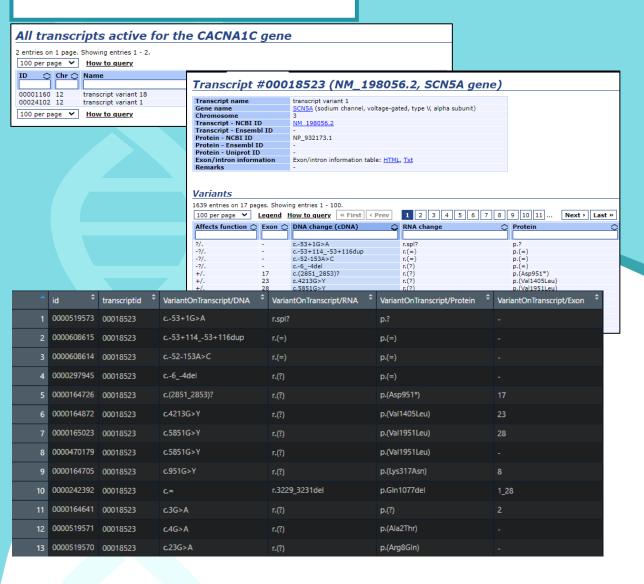


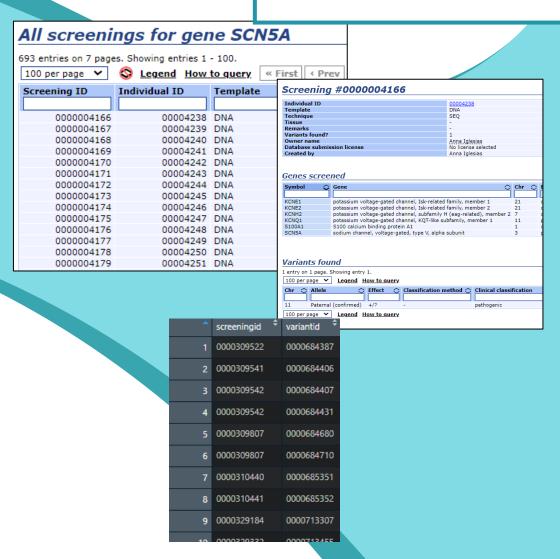
EXTENSION

SILE

VARIANTS ON TRANSCRIPTS

SCREENINGS TO VARIANTS





with the individuals list

Takes Disease ID from each "href" and associate it

INDIVIDUALS TO DISEASES

11 00047339

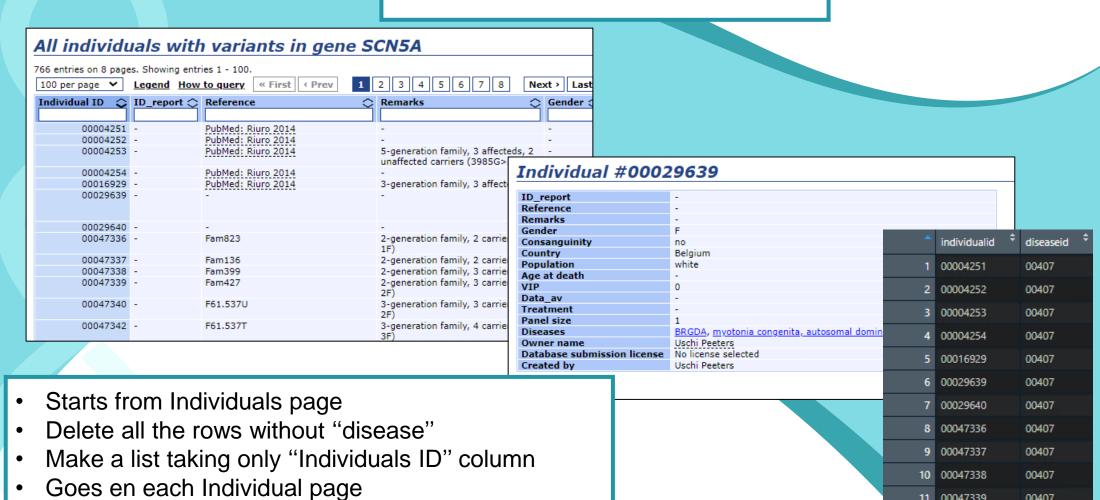
12 00047340

13 00047342

00407

00407

00407



disease"

PHENOTYPE

Inheritan

Familial

Phenotype/Age

Isolated (sporadic)

Isolated (sporadic)

Familial

Familial

Familial

Familial

Familial

Unknown

Familial, autosomal dominant

Familial, autosomal dominant

Phenotype/Additional

Diagnosis/Definite

Phenotype/Inheritance

0000035751 00407

0000034668 00407

0000035753 00407

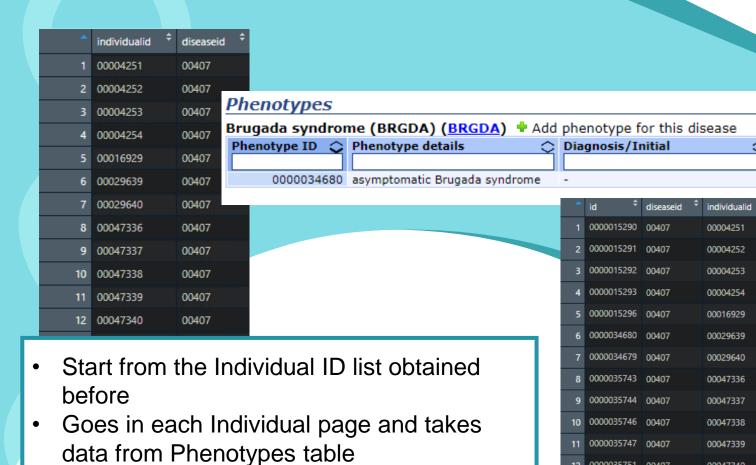
0000034743 00407

00047340

00047342

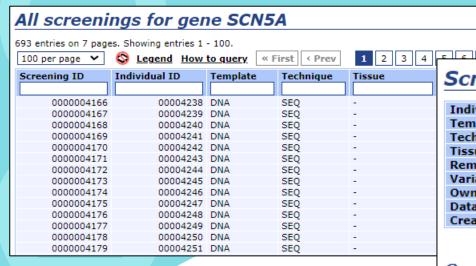
00047343

00047469

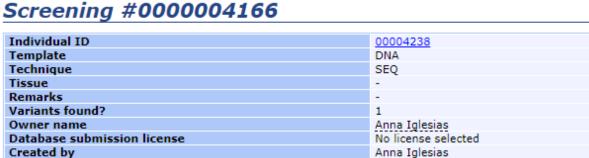


Merges the obtained db with "individual to

SCREENINGS TO GENES



- Start from the Screenings table
- Goes in each Screening page and takes data from Genes Screened table
- Keeps only Screening ID and Gene ID



Genes screened

Symbol (Gene	Chr	\bigcirc Ba	and	Transci	
					A	_
KCNE1	potassium voltage-gated channel, Isk-related family, member 1	21		screeningic	d [∓] genei	id T
KCNE2	potassium voltage-gated channel, Isk-related family, member 2	21		1 000000416	66 KCNE	E1
KCNH2	potassium voltage-gated channel, subfamily H (eag-related), member 2	7		200000440		_
KCNQ1	potassium voltage-gated channel, KQT-like subfamily, member 1	11		2 000000416	66 KCNE	2
S100A1	S100 calcium binding protein A1	1		3 000000416	6 KCNH	1 2
SCN5A	sodium channel, voltage-gated, type V, alpha subunit	3				
				4 000000416	66 KCNC	21
				5 000000416	56 \$100	A1
				6 000000416	6 SCN5	5A
					5 50113	
				7 000000416	7 KCNE	E1
				8 000000416	7 KCNE	= 2
				9 000000416	7 KCNH	12
				10 000000416	7 KCNC	Q1

VARIANTS ON GENOME

VariantOnGenome/DNA

g.95372567C>T

g.95372567C>T

q.95372567C>T

g.95372567C>T

g.95372685C>T

g.95372734G>A

q.95372734G>A

g.95372734G>T

g.95372734G>T

q.95372764C>T

g.95372786C>T

PDE6C 000034

PDE6C_000002

PDE6C 000002

PDE6C_000001

PDE6C 000001

PDE6C 000025

PDE6C_000084

02330

02325

02327

02330

01943

02327

00000

g.95372571 95372581del

g.(? 95372482) (95372963

Variant #0000519573 (NC_000003.11:g.38691021C>T,

7 0000305372 3

8 0000305372 3

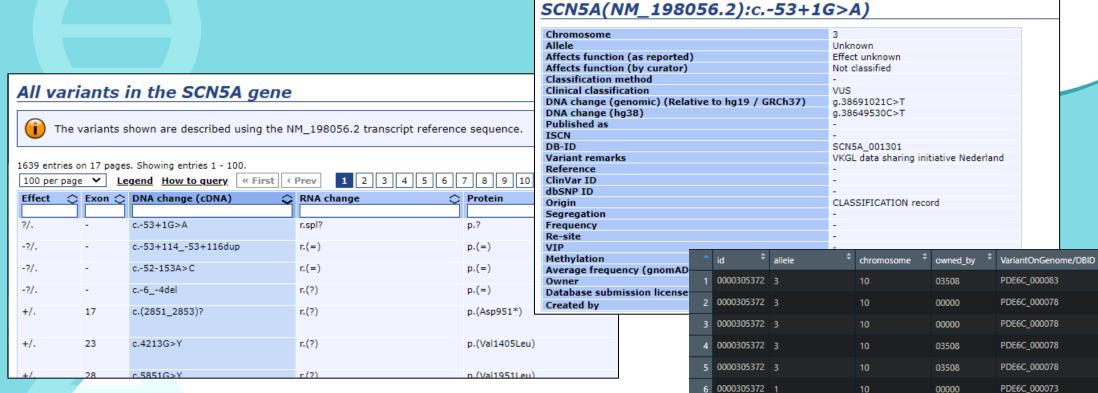
9 0000305372 3

10 0000305372 3

11 0000305372 3

12 0000305372 3

13 0000305372 Both (homozygous) 10

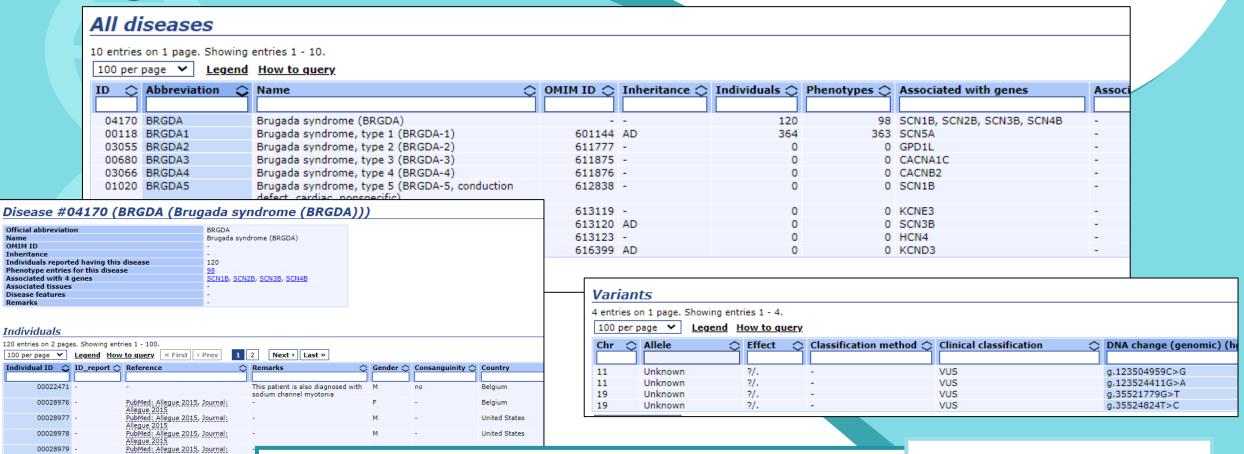


- Takes the whole variants table and also put the href in each row in a separate list
- Goes in each Variant page and takes data from the table
- Merges the two tables

PubMed: Allegue 2015, Journal:

00028980

RESEARCH BY PHENOTYPE



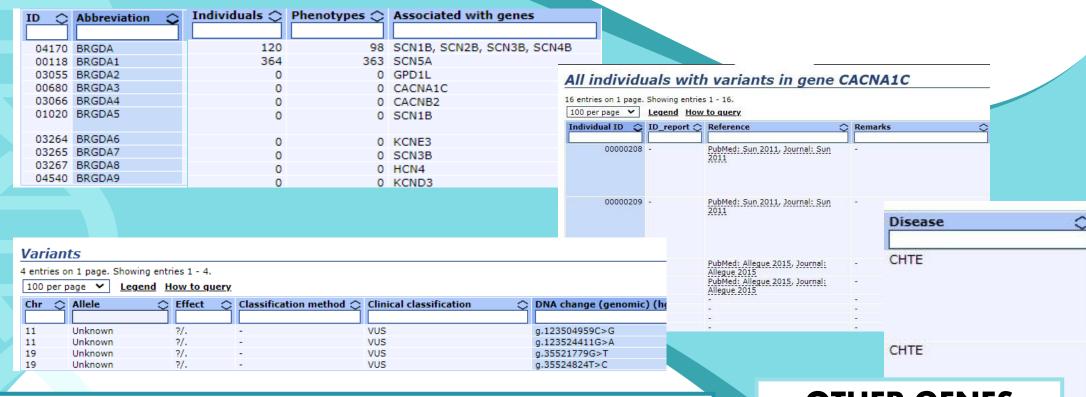
• Starts from the Disease page related to the Disease choosen

Goes into each row and takes data from each individual table page

- Goes into each Individual page and takes data from variants table
- Merges Individuals and Variants tables

BRGDA & BRGDA1

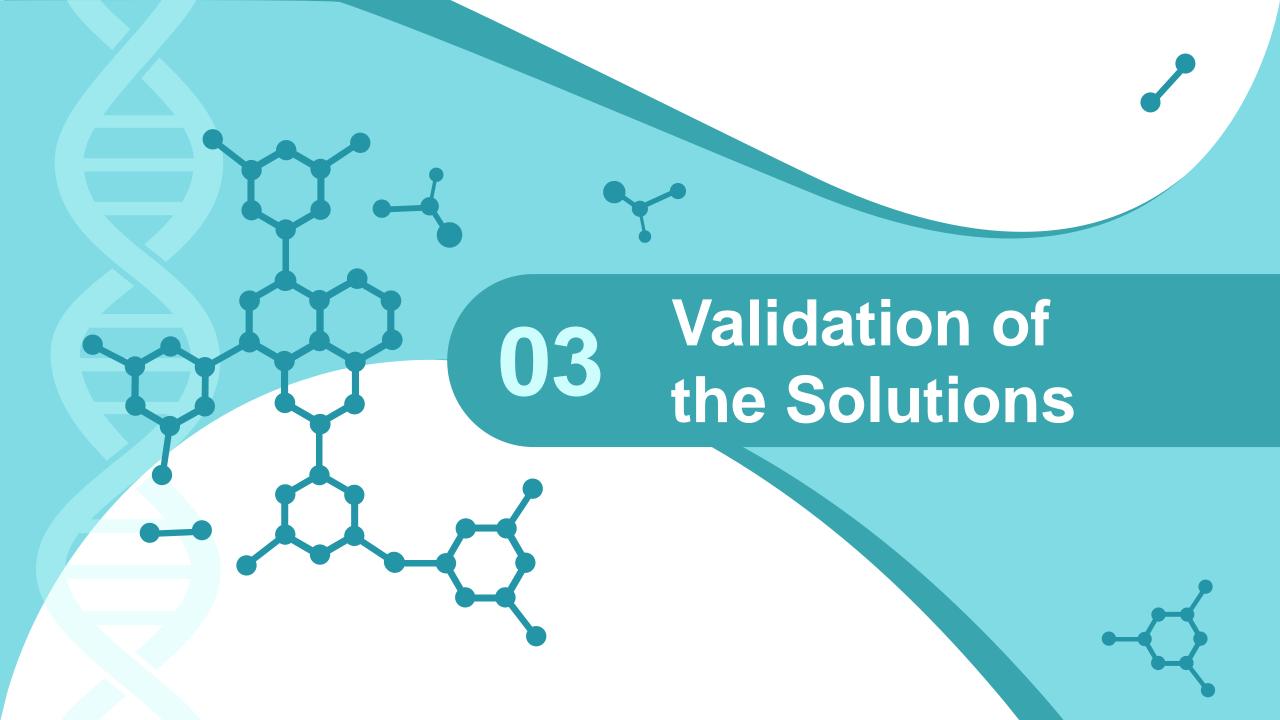
RESEARCH BY PHENOTYPE



- Takes the list of the Associated Genes from each row who has Individuals = 0
- Goes in each Individuals page associated with each gene in the list
- Goes into each Individual page and takes data from variants table
- Merges Individuals and Variants tables
- Filter the obtained table by the column "Disease" by taking only the related ones

OTHER GENES

BRGDA
BRGDA
SUD
BRGDA
BRGDA
SUD



BRGDA & BRGDA1

BRGDA:

From 120 individuals, 420 variations data

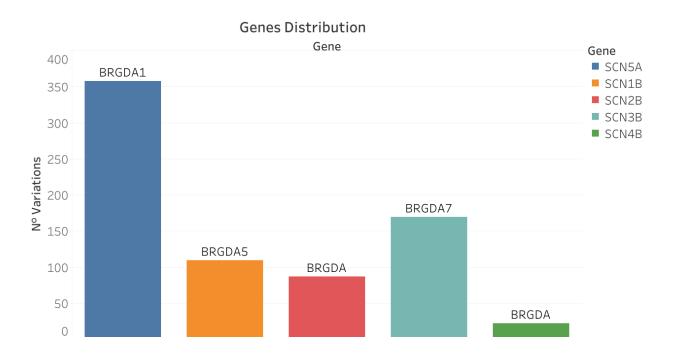
Study of *Peeters* on 74 unrelated individuals

[SCN1B, SCN2B, SCN3B, SCN4B]

BRGDA 1:

From 364 individuals, 366 variations data

Genes Distribution



BRGDA & BRGDA1

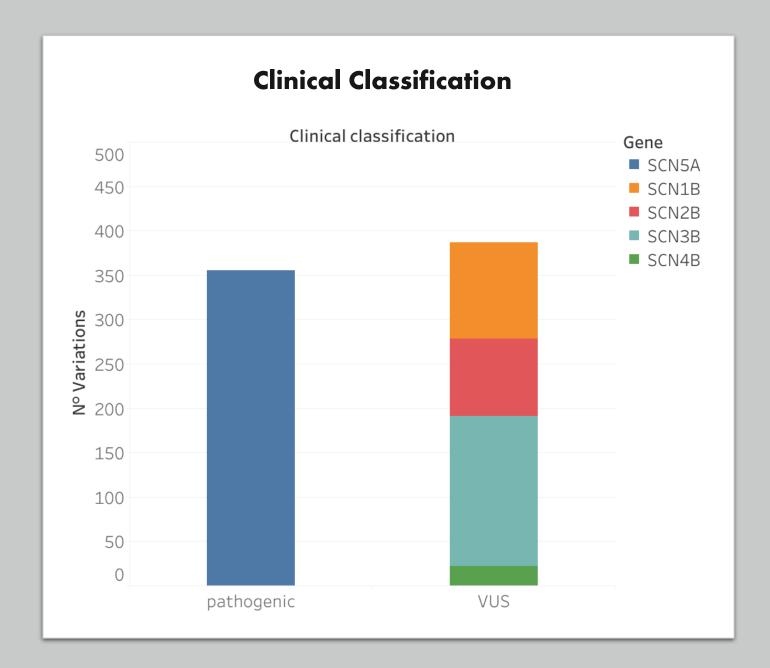
BRGDA:

Totality of observations classified as "VUS"

(Variant of Uncertain Significance)

BRGDA 1:

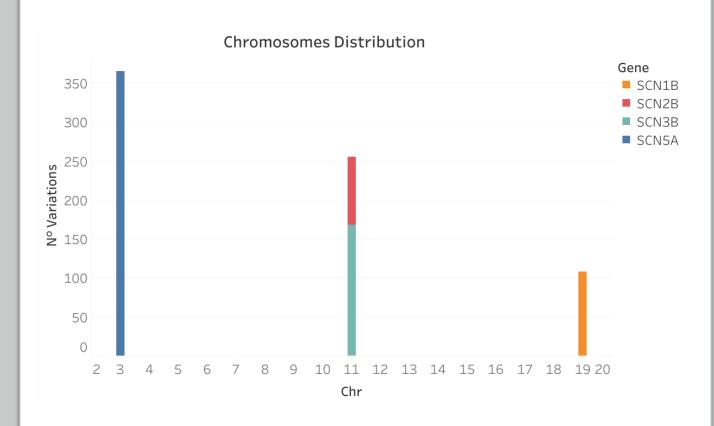
Totality of observations classified as "pathogenic"



BRGDA & BRGDA1

- SCN5A, typically associated with BRGDA1, mutation in **chromosome 3**
- SCN1B, typically associated with BRGDA5, mutations in chromosome 19
- SCN3B, typically associated with BRGDA7, mutations in **chromosome 11**

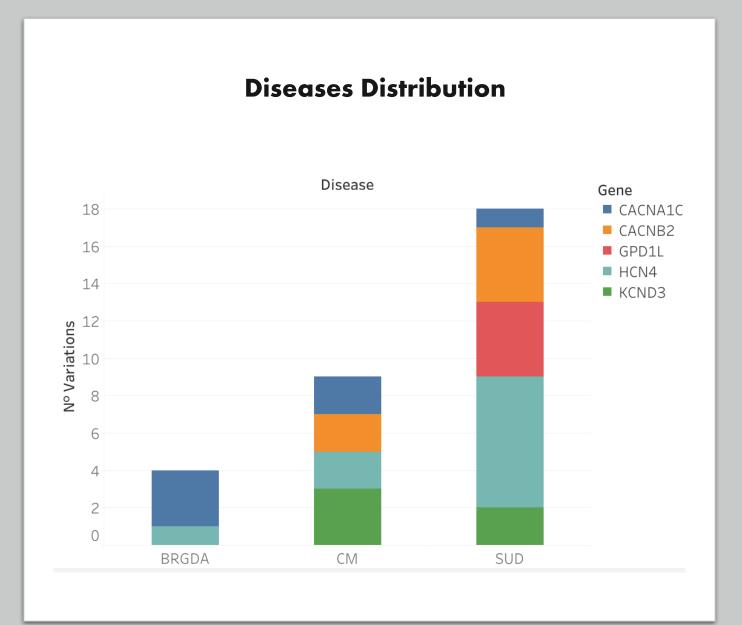
Chromosomes Distribution



From the **921** variations obtained:

- Deleted the ones already present in the previous database by "individual ID"
- Selecting the ones labeled as:
 - BRGDA
 - CM (Cardiomyopathy), associated with strong arrhythmias
 - SUD (Sudden Unexplained Death), synonym of Brugada in MedGen

Resulting db composed by 31 variations



Variations in:

GPD1L (4)

CACNA1C (6)

CACNB (6)

HCN4 (10)

KCND3 (5)

Typically associated with:

BRGDA2

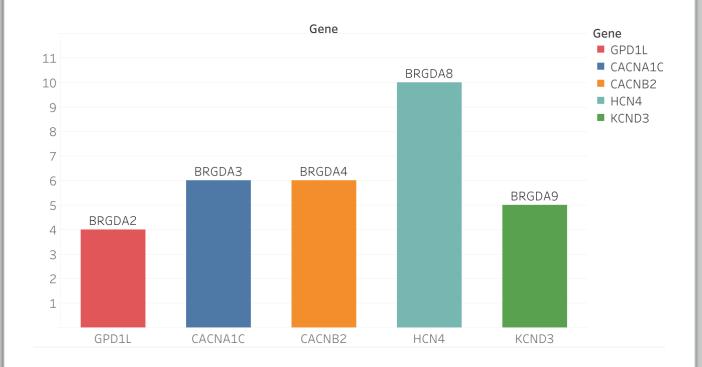
BRGDA3

BRGDA4

BRGDA8

BRGDA9

Genes Distribution



15 variations labeled as "VUS"

GPD1L labeled as "pathogenic" in all observations

KCND3 in the majority "likely pathogenic"



pathogenic

VUS

0

likely benign

likely pathogenic

GPD1L, mutation in **chromosome 3** (Brs2)

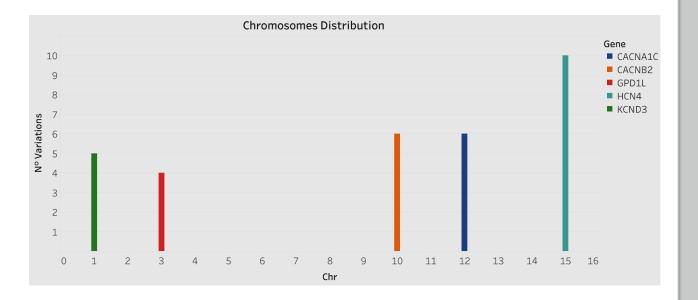
CACNA1C, mutation in **chromosome 12** (BrS3)

CACNB2, mutation in chromosome 1 (BrS4)

HCN4, mutation in **chromosome 15** (Brs8)

KCND3, mutation in **chromosome 1** (BrS9)

Chromosomes Distribution



DISCUSSION

DOWNLOAD:

Download all this gene's data	Download all data
Notes	Establishment of this gene
	supported by the <u>Leiden Ur</u>
	(LUMC), Leiden, Nederland
Date created	May 03, 2013
Date last updated	September 17, 2021

PRO:

Reduce time to collect data

CONS:

- Feature available for only a small group of the genes in LOVD
- **Lower quality** of data obtained:
 - Data in download file are updated manually, so most of the time the information present in it are different from the ones on the website
- Can't obtain the information present in "Research by Phenotype", fundamental for the analysis carried out by PROS

DISCUSSION

TIME:

Library "*rvest*" only obtains data in the first table (**max 100**), even manually changing the url the result doesn't change

Usually variants, screenings and individuals have more than 100 entries

Library "Rselenium" overcomes the problem by simulating the navigation and clicking the button to take data from the subsequent tables

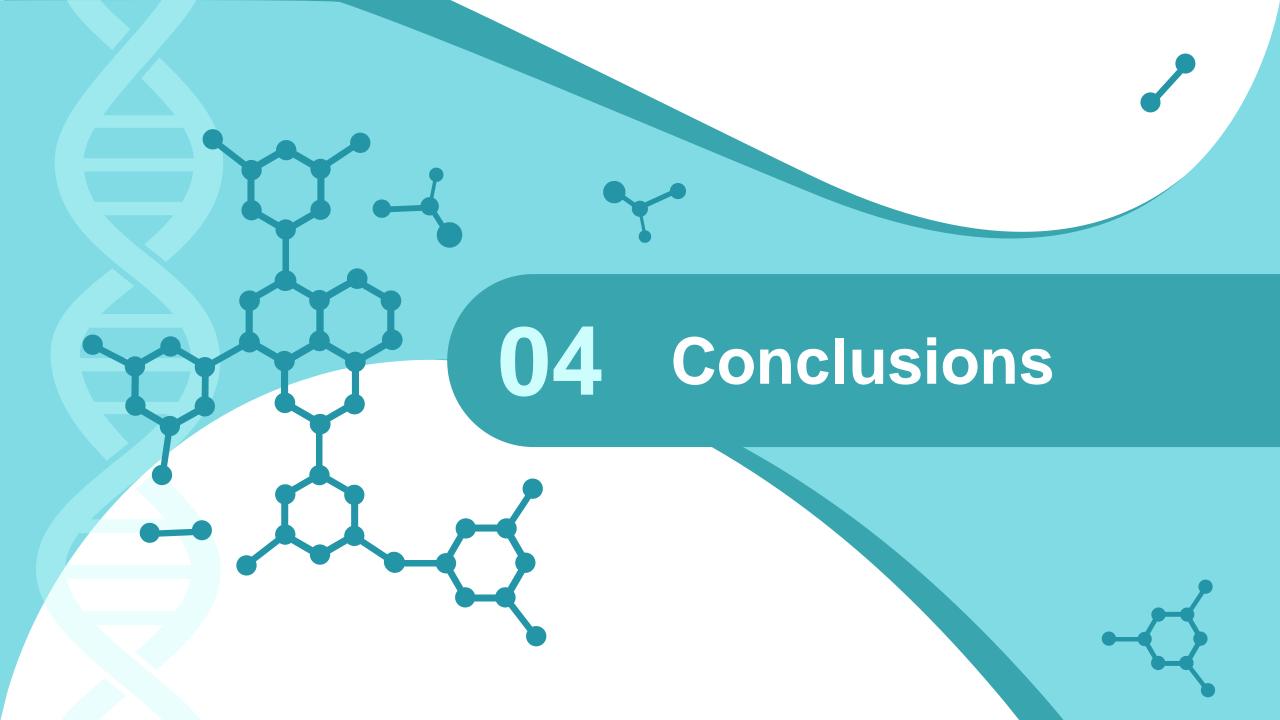
Screenings page, even in normal navigation, takes more time than others to load

No API available, so web scraping is the only way to obtain data for the moment

DATA:

Brugada Syndrome is a **rare** disease, the most common type is BRGDA1

Types from **BRGDA2** to **BRGDA9** are even **more rare** so only few data are available and in most of the cases are labeled with the generic name "**BRGDA**"



CONCLUSIONS

OBJECTIVES:

- 1) After applying SILE method to Brugada Syndrome, noticed a reduced presence of data
- 2) For this reason, another source of data has been **added** to the ones already present: the biological database **LOVD**
- **3) Evaluated** the effectiveness of the extension of the research module by **applying** it to analysis of **Brugada** Syndome

CONCLUSIONS

BRGDA1:

Most of the data obtained (45%) refers to the variations of SCN5A as "pathological", which is associated with BRGDA1 (most widespread type)

BRGDA:

Most of the data belonging to **BRGDA** has variations in genes SCN1B and SCN3B, generally associated to BRGDA5 and BRGDA7

Clinical classification labeled as "VUS" (not yet possible to determine if pathological or benign)

OTHERS:

Variations of GPD1L, CACNA1C, CACNB2, HCN4, KCND3, labeled in LOVD as "CM" and "SUD"

Data keeped because **generally associated** with BRGDA2, BRGDA3, BRGDA4, BRGDA8, BRGDA9

More accurate assessment in **following steps** and evaluation by **panel of experts** to take this decision and to evaluate the overall **quality** of acquired data

FUTURE WORKS

Algorithm developed able to add biological data **not only** related to **Brugada**, that **overcomes** the **limitations** imposed by the database, allowing to search data starting from a selected **phenotype**

Platform in future will become more **user friendly**, in order to make it usable by professionals in field of Precision Medicine

Since biological data are **scattered** and **heterogeneous**, adding more **data sources** could help to perform a better analysis as the cost of more time for searching and standardizing data

Cyclical update in automated way of the data from the various data sources

Perform researches starting from a specific Gene

Follow further developments in the data currently available, especially **VUS**