Rare disorder diagnostics

From whole exome sequencing to regulatory regions in whole genomes.

Paweł Sztromwasser
Department of Biostatistics and Translational Medicine
Medical University of Lodz

SMM, Kraków
12/11/2018

NATIONAL SCIENCE CENTRE
POLAND

Outline

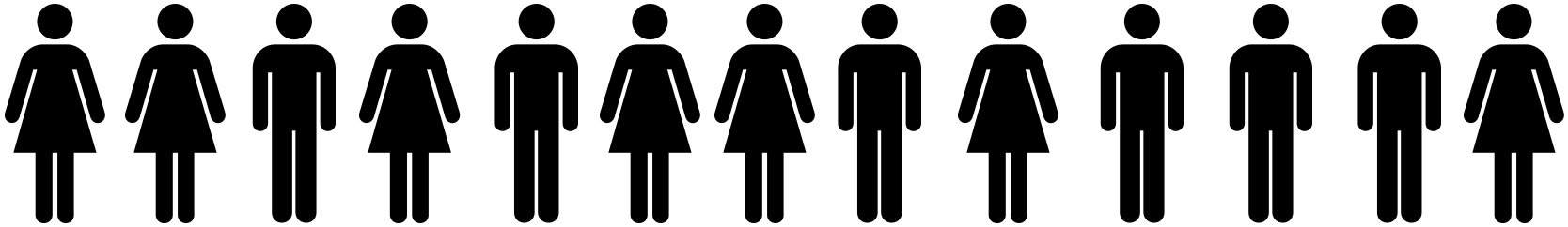
1. NGS in medical diagnostics - a success story
2. Challenges
3. Remus - REgulatory MUtation Search
4. Case study - RCAD

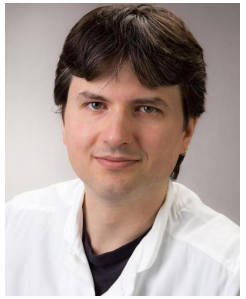
BTM

Haukeland University Hospital

University of Bergen

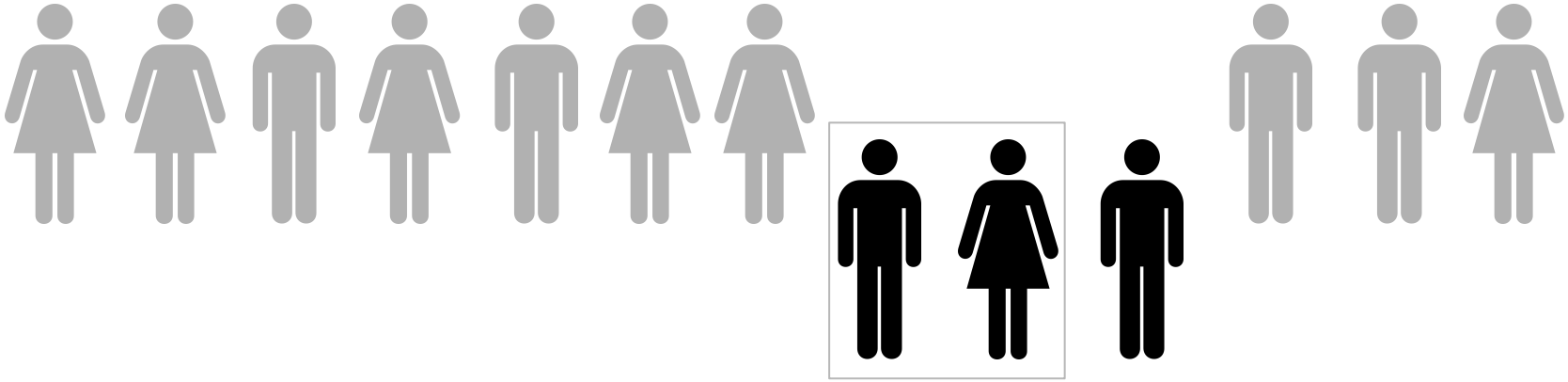
Hereditary neurological disorders





Haris Tzoulis

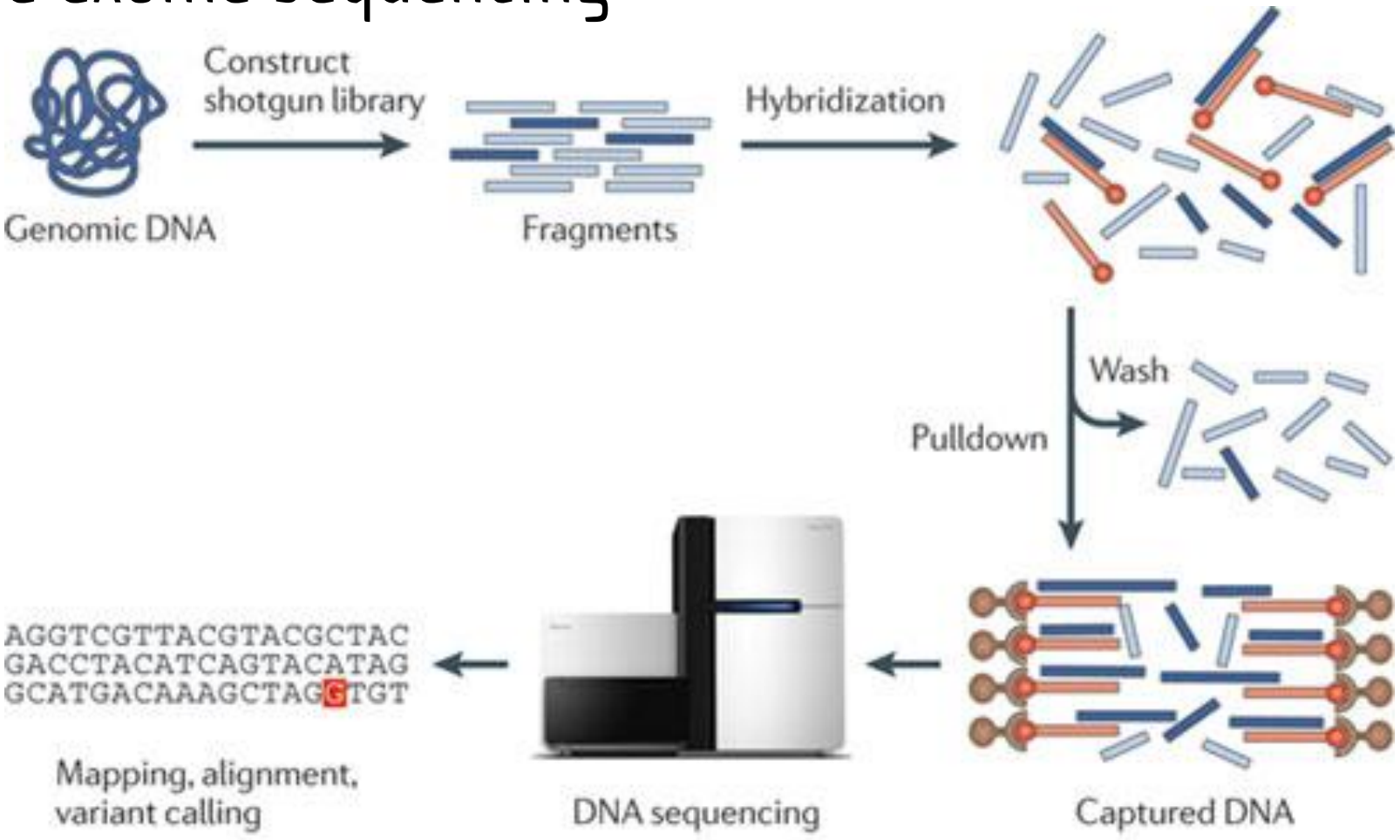
Hereditary neurological disorders



progressive encephalopathy diagnosed at age 1
dystonia, epilepsy, episodic exacerbations

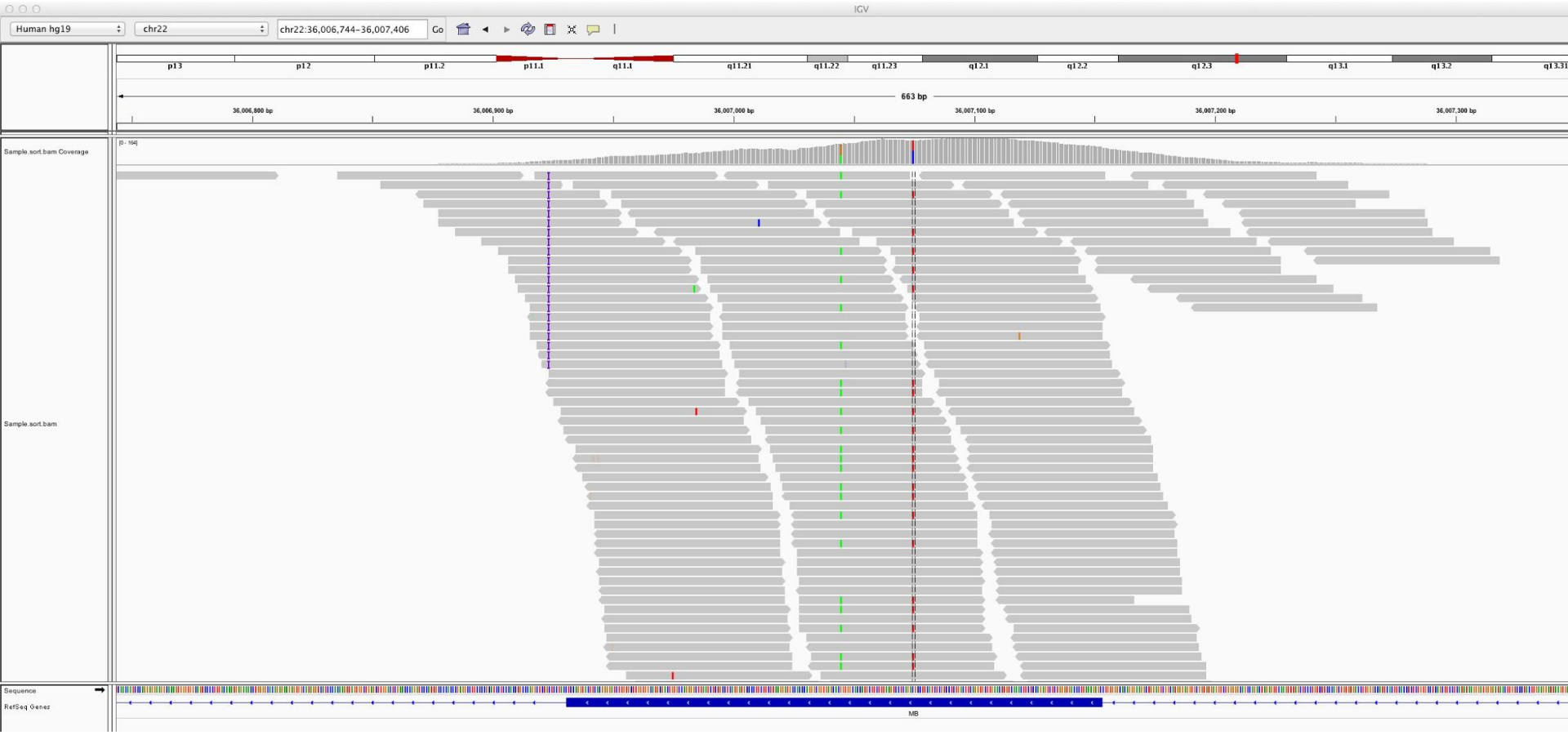
parents unaffected

Whole exome sequencing

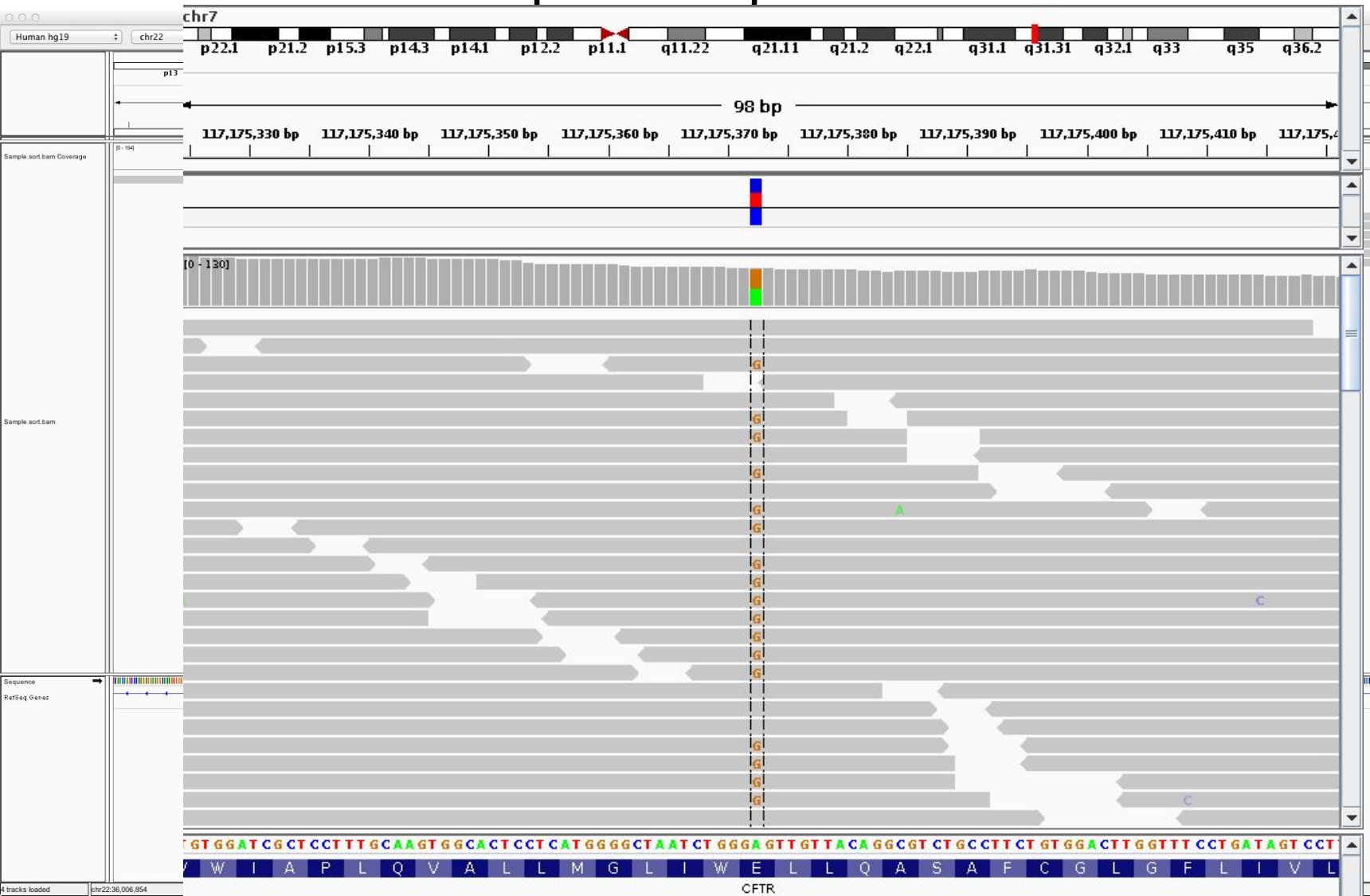


Bamshad MJ, et al (2011). Exome sequencing as a tool for Mendelian disease gene discovery, Nature Reviews Genetics 12, 745-755

Whole exome sequencing

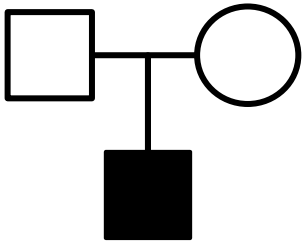


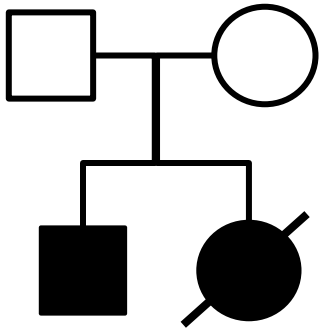
Whole exome sequencing



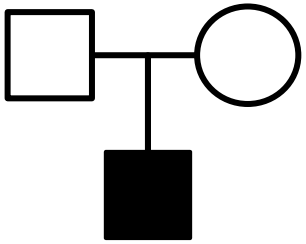
Dystonia genes

Symbol	OMIM	Gene	Locus	Alt Name
DYT1	128100	TOR1A	9q34	Early-onset torsion dystonia
DYT2	224500	HPCA	1p35-p34.2	Autosomal recessive primary isolated dystonia
DYT3	314250	TAF1	Xq13	X-linked dystonia-parkinsonism
DYT4	128101	TUBB4[4]	19p13.12-13	Autosomal dominant whispering dysphonia
DYT5a	128230	GCH1	14q22.1-q22.2	Autosomal dominant dopamine-responsive dystonia
DYT5b	191290	TH	11p15.5	Autosomal recessive dopamine-responsive dystonia
DYT6	602629	THAP1	8p11.21	Autosomal dominant dystonia with cranio-cervical predilection
DYT7	602124	unknown	18p (questionable)	Autosomal dominant primary focal cervical dystonia
DYT8	118800	MR1	2q35	Paroxysmal nonkinesigenic dyskinesia
DYT9	601042	SLC2A1	1p35-p31.3	Episodic choreoathetosis/spasticity (now known to be synonymous with DYT18)
DYT10	128200	PRRT2	16p11.2-q12.1	Paroxysmal kinesigenic dyskinesia
DYT11	159900	SGCE	7q21	Myoclonic dystonia
DYT12	128235	ATP1A3	19q12-q13.2	Rapid onset dystonia parkinsonism and alternating hemiplegia of childhood
DYT13	607671	unknown, near D1S2667[5]	1p36.32-p36.13	Autosomal dominant cranio-cervical/upper limb dystonia in one Italian family
DYT14	See DYT5			
DYT15	607488	unknown	18p11[6]	Myoclonic dystonia not linked to SGCE mutations
DYT16	612067	PRKRA	2q31.3	Autosomal recessive young onset dystonia parkinsonism
DYT17	612406	unknown, near D20S107[7]	20p11.2-q13.12	Autosomal recessive dystonia in one family
DYT18	612126	SLC2A1	1p35-p31.3	Paroxysmal exercise-induced dyskinesia
DYT19	611031	probably PRRT2	16q13-q22.1	Episodic kinesigenic dyskinesia 2, probably synonymous with DYT10
DYT20	611147	unknown	2q31	Paroxysmal nonkinesigenic dyskinesia 2
DYT21	614588	unknown	2q14.3-q21.3	Late-onset torsion dystonia
DYT24	610110	ANO3[8]	11p14.2	Autosomal dominant cranio-cervical dystonia with prominent tremor





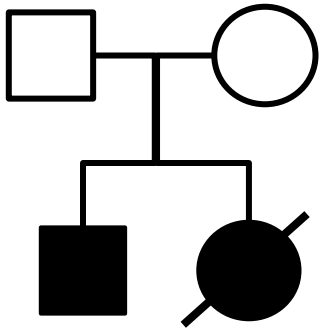
SLC19A3



SLC19A3:

c.337T>C, p.Y113H

c.541T>C, p.S181P



SLC19A3:

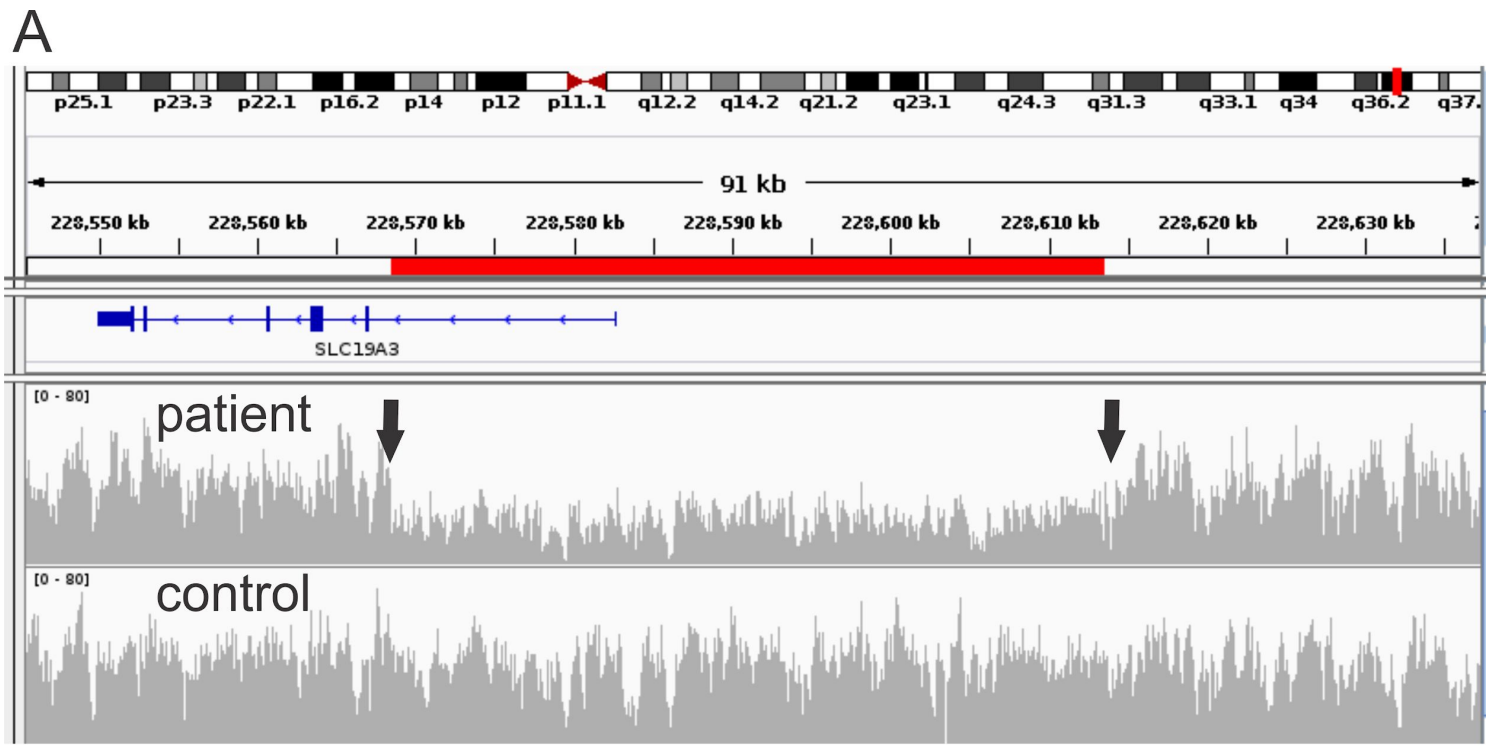
c.337T>C, p.Y113H

???

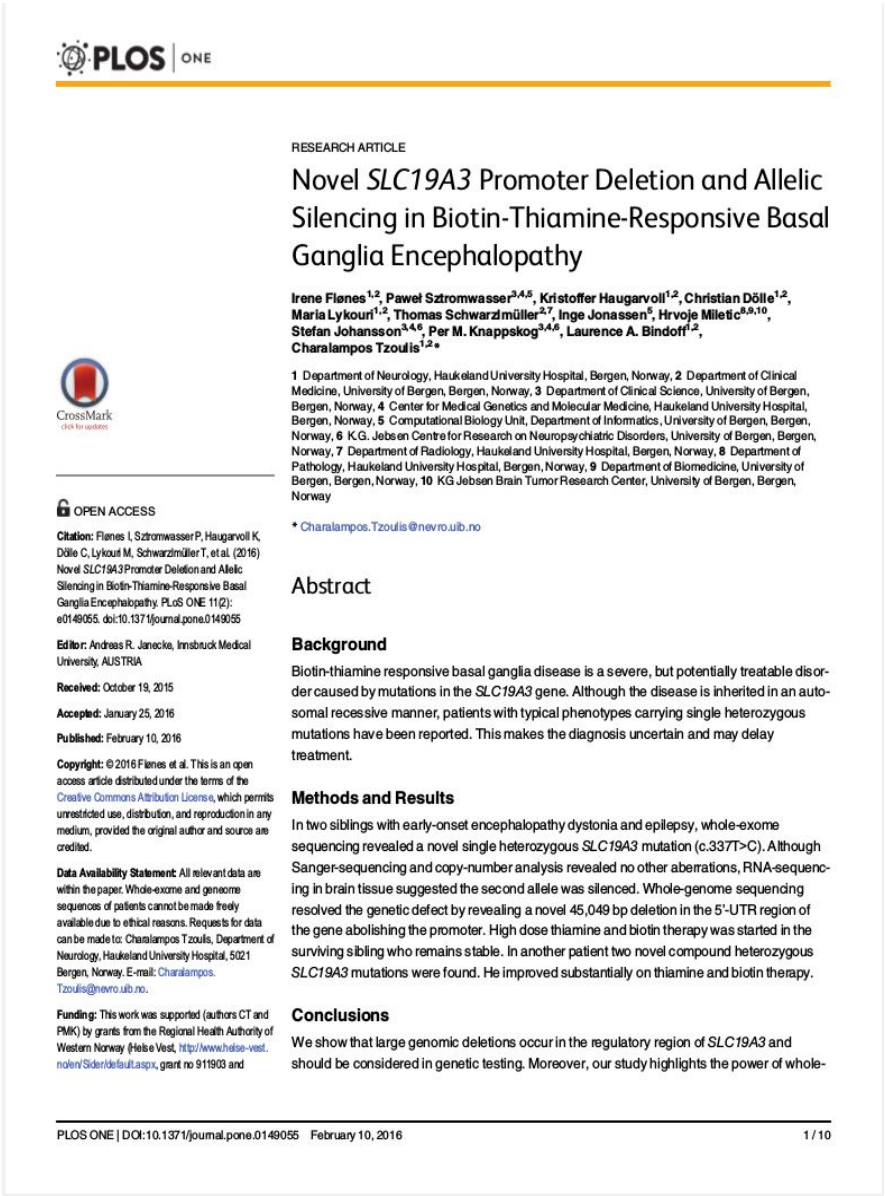
RNA sequencing

in brain, only the mutated allele was expressed...

Whole-genome sequencing



I.Flønes, et al. Novel SLC19A3 Promoter Deletion and Allelic Silencing in Biotin-Thiamine-Responsive Basal Ganglia Encephalopathy. PLoS ONE, 2016



PLOS | ONE

RESEARCH ARTICLE

Novel *SLC19A3* Promoter Deletion and Allelic Silencing in Biotin-Thiamine-Responsive Basal Ganglia Encephalopathy

Irene Fløtnes[1,2], Paweł Sztromwasser[3,4,5], Kristoffer Haugarvoll[1,2], Christian Dölle[1,2], Maria Lykouri[1,2], Thomas Schwarzlmüller[2,7], Inge Jonassen[5], Hrvoje Miletic[8,9,10], Stefan Johansson[3,4,6], Per M. Knappskog[3,4,6], Laurence A. Bindoff[1,2], Charalampos Tzoulis[1,2]*

1 Department of Neurology, Haukeland University Hospital, Bergen, Norway, **2** Department of Clinical Medicine, University of Bergen, Bergen, Norway, **3** Department of Clinical Science, University of Bergen, Bergen, Norway, **4** Center for Medical Genetics and Molecular Medicine, Haukeland University Hospital, Bergen, Norway, **5** Computational Biology Unit, Department of Informatics, University of Bergen, Bergen, Norway, **6** K.G. Jebsen Centre for Research on Neuropsychiatric Disorders, University of Bergen, Bergen, Norway, **7** Department of Radiology, Haukeland University Hospital, Bergen, Norway, **8** Department of Pathology, Haukeland University Hospital, Bergen, Norway, **9** Department of Biomedicine, University of Bergen, Bergen, Norway, **10** KG Jebsen Brain Tumor Research Center, University of Bergen, Bergen, Norway

* Charalampos.Tzoulis@nevro.uib.no

OPEN ACCESS

Citation: Fløtnes I, Sztromwasser P, Haugarvoll K, Dölle C, Lykouri M, Schwarzlmüller T, et al. (2016) Novel *SLC19A3* Promoter Deletion and Allelic Silencing in Biotin-Thiamine-Responsive Basal Ganglia Encephalopathy. PLoS ONE 11(2): e0149055. doi:10.1371/journal.pone.0149055

Editor: Andreas R. Janecke, Innsbruck Medical University, AUSTRIA

Received: October 19, 2015

Accepted: January 25, 2016

Published: February 10, 2016

Data Availability Statement: All relevant data are within the paper. Whole-exome and genome sequences of patients cannot be made freely available due to ethical reasons. Requests for data can be made to: Charalampos Tzoulis, Department of Neurology, Haukeland University Hospital, 5021 Bergen, Norway. E-mail: Charalampos.Tzoulis@nevro.uib.no.

Funding: This work was supported (authors CT and PMK) by grants from the Regional Health Authority of Western Norway (Helse Vest, http://www.helse-vest.no/en/Sider/default.aspx, grant no 911903 and

Abstract

Background

Biotin-thiamine responsive basal ganglia disease is a severe, but potentially treatable disorder caused by mutations in the *SLC19A3* gene. Although the disease is inherited in an autosomal recessive manner, patients with typical phenotypes carrying single heterozygous mutations have been reported. This makes the diagnosis uncertain and may delay treatment.

Methods and Results

In two siblings with early-onset encephalopathy dystonia and epilepsy, whole-exome sequencing revealed a novel single heterozygous *SLC19A3* mutation (c.337T>C). Although Sanger-sequencing and copy-number analysis revealed no other aberrations, RNA-sequencing in brain tissue suggested the second allele was silenced. Whole-genome sequencing resolved the genetic defect by revealing a novel 45,049 bp deletion in the 5'-UTR region of the gene abolishing the promoter. High dose thiamine and biotin therapy was started in the surviving sibling who remains stable. In another patient two novel compound heterozygous *SLC19A3* mutations were found. He improved substantially on thiamine and biotin therapy.

Conclusions

We show that large genomic deletions occur in the regulatory region of *SLC19A3* and should be considered in genetic testing. Moreover, our study highlights the power of whole-

I.Fløtnes, et al. Novel SLC19A3 Promoter Deletion and Allelic Silencing in Biotin-Thiamine-Responsive Basal Ganglia Encephalopathy. PLoS ONE, 2016

BTM

MEDICAL UNIVERSITY OF LODZ

rgens Tidende
NR. 73 UKE 11 - 2015 - 148. ÅRGANG
LØSSALG KR 35

Gentest på sykehus ble redningen for Frank (23)

GOD VENN: Frank S. Arnesen (23) har en sjelden genfeil. En test ga svaret. Her er han i sofaen sammen med hunden Sharky.
FOTO: HÅVARD BJELLAND

Legene fant ikke ut hva Frank Stræng Arnesen feilte. Han var nær ved å dø av den mystiske sykdommen. Så tok han en gentest ved Haukeland universitetssykehus. Det ble et vendepunkt for 23-åringen. Nå er han på bedringens vei.

NYHETER // SIDE 4–6

Acknowledgement

Irene Fløne s

Charalampos Tzoulis

Stefan Johansson

Vidar Steen

I.Fløne s, et al. Novel SLC19A3 Promoter Deletion and Allelic Silencing in Biotin-Thiamine-Responsive Basal Ganglia Encephalopathy. PLoS ONE, 2016

Challenges

WES - 25-50 % success rate (diagnostic yield)

WGS - 50 % success rate

70-90% of the genome accessible to short read technologies

Structural variants - small number but equal in extent

Gilissen et al. 2014; Yang et al. 2013; Lee et al. 2015; Valencia et al. 2015; Lee and Schatz, 2012; The 1000 Genomes Consortium, 2015

Challenges

WES - 25-50 % success rate (diagnostic yield)

WGS - 50 % success rate

70-90% of the genome accessible to short read technologies

Structural variants - small number but equal in extent

Focused on protein coding – 1% of the genome

Gilissen et al. 2014; Yang et al. 2013; Lee et al. 2015; Valencia et al. 2015; Lee and Schatz, 2012; The 1000 Genomes Consortium, 2015

Non-coding genome

99% of the genome is non-coding

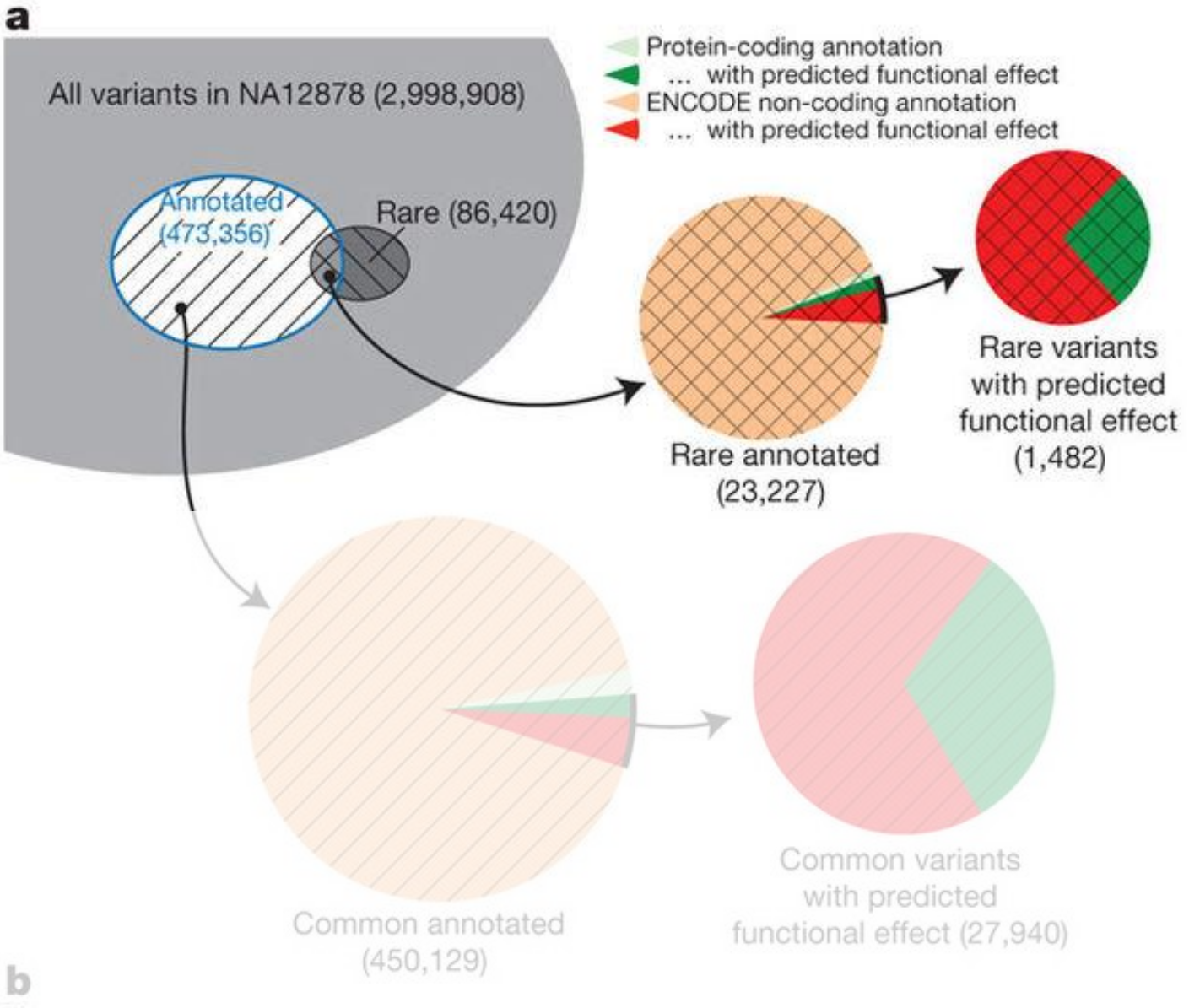
80.4% of the genome could be assigned a function

Enhancer-like - 400,000 (**8%** of the genome), promoter like - 70,000

8.8k small RNAs, **9.6k** long non-coding RNAs

The ENCODE Project Consortium, 2012

Challenges



The ENCODE Project Consortium, 2012

453

cases of **noncoding**, **regulatory** mutations causing a rare monogenic disease, according to **OMIM** database

Smedley et al. , 2016

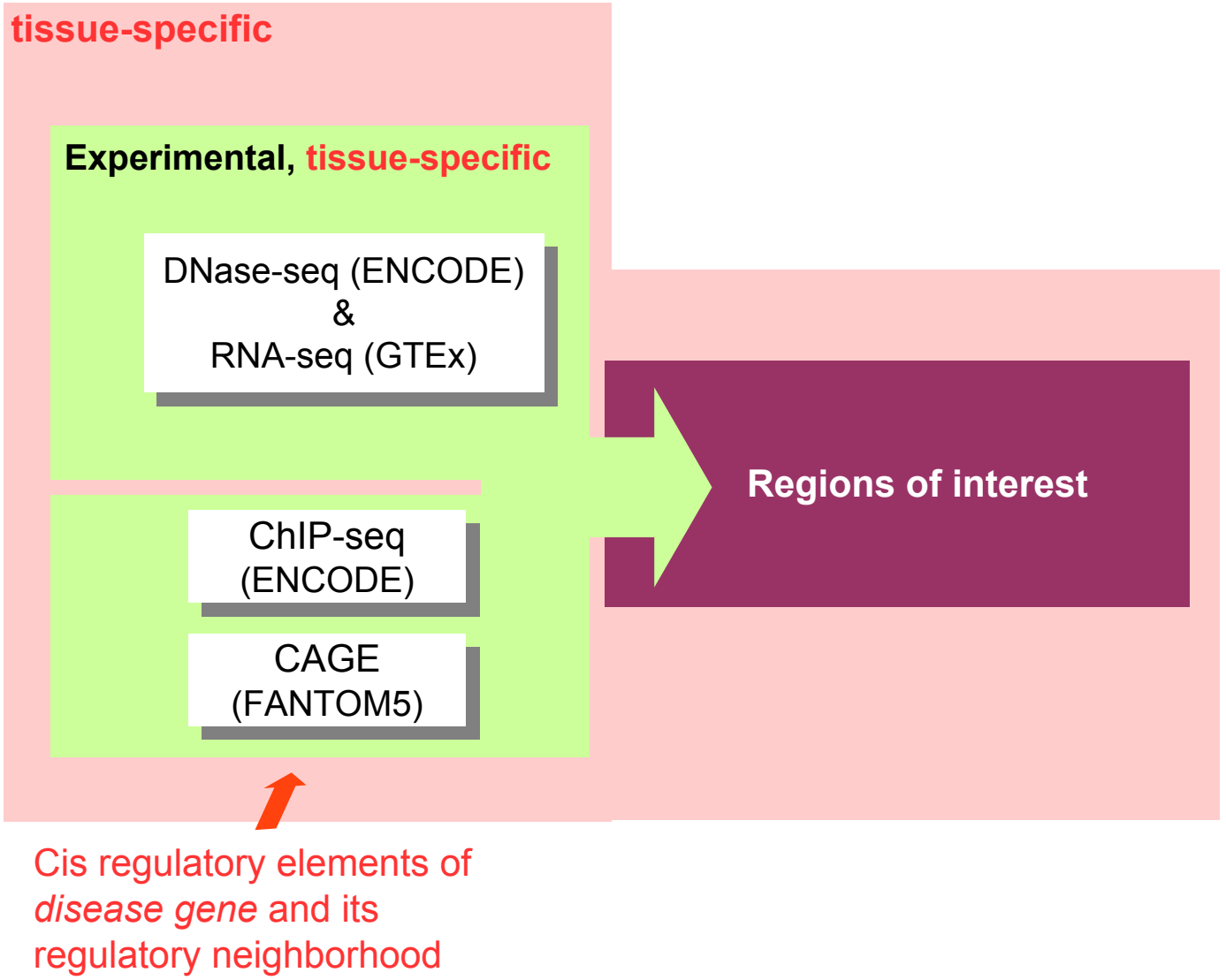
Pathogenic noncoding mutations

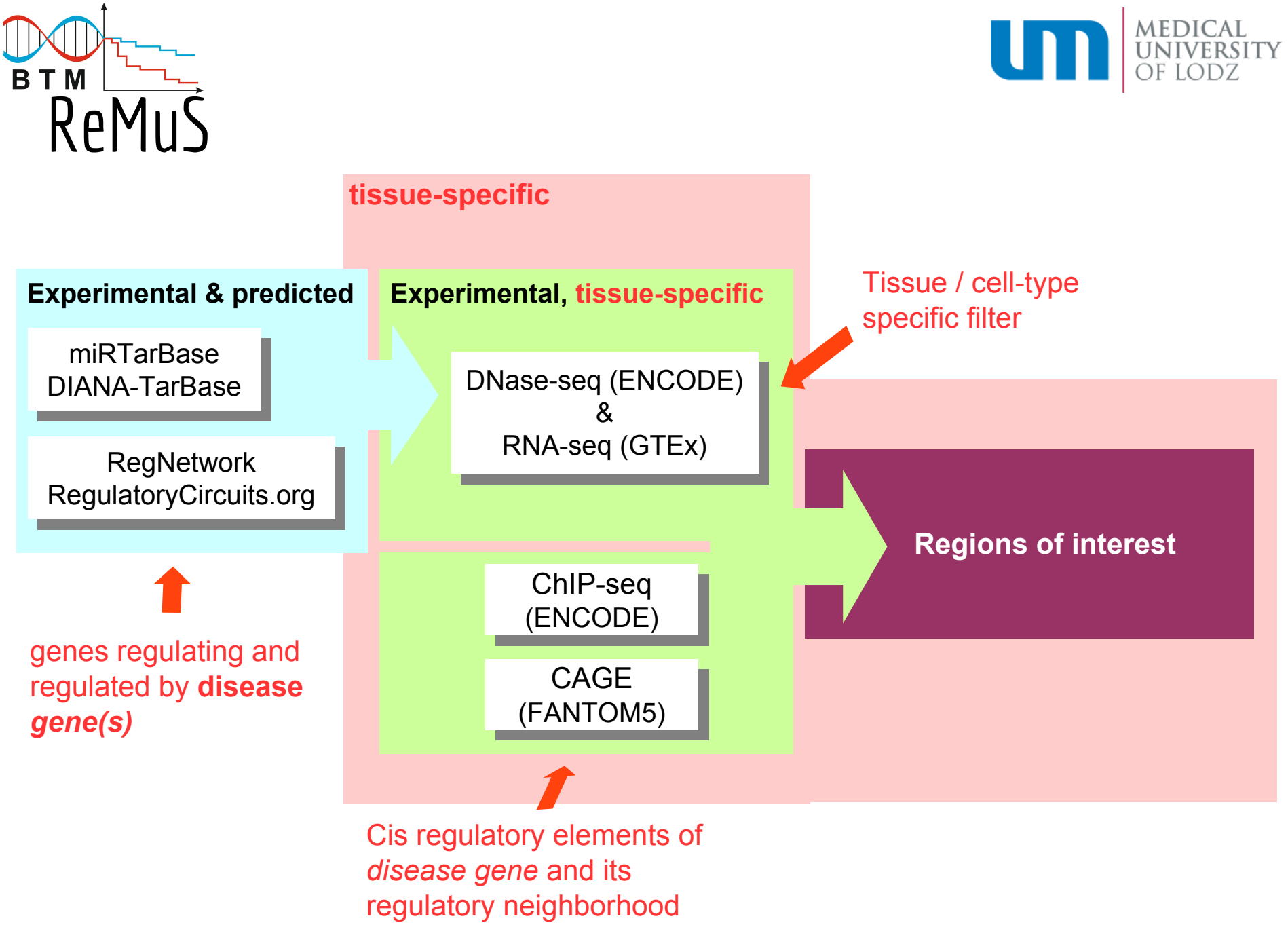
Table 1. Mendelian Regulatory Mutations

Category	Example	Count
Enhancer	triphalangeal thumb, type I (*SHH* [MIM: 174500])	42
Promoter	hemophilia B (*F9* [MIM: 306900])	142
5′ UTR		153
Transcription (core promoter)	acute intermittent porphyria (*HMBS* [MIM: 176000])	52/153
uORF	Marie Unna hereditary hypotrichosis (*HR* [MIM: 146550])	37/153
Secondary structure	hyperferritemia cataract syndrome (*FTL* [MIM: 600886])	31/153
Kozak sequence	beta thalassemia (*HBB* [MIM: 613985])	2/153
Unclassified	thrombocytopenia 2 (*ANKRD26* [MIM: 188000])	31/153
3′ UTR		43
Polyadenylation	permanent neonatal diabetes (*INS* [MIM: 606176])	14/43
miRNA binding	autosomal-dominant spastic paraplegia 31 (*REEP1* [MIM: 610250])	5/43
Other	autosomal-dominant myopia 21 (*ZNF644* [MIM: 614167])	24/43
Large non-coding RNA gene	microcephalic osteodysplastic primordial dwarfism, type 1 (*RNU4ATAC* [MIM: 210710])	65
MicroRNA gene	autosomal-dominant deafness 50 (*MIR96* [MIM: 613074])	5
Imprinting control region	Beckwith-Wiedemann syndrome (*H19* [MIM: 130650])	3
Total		453
Total single-nucleotide variants		406

Smedley et al. , 2016

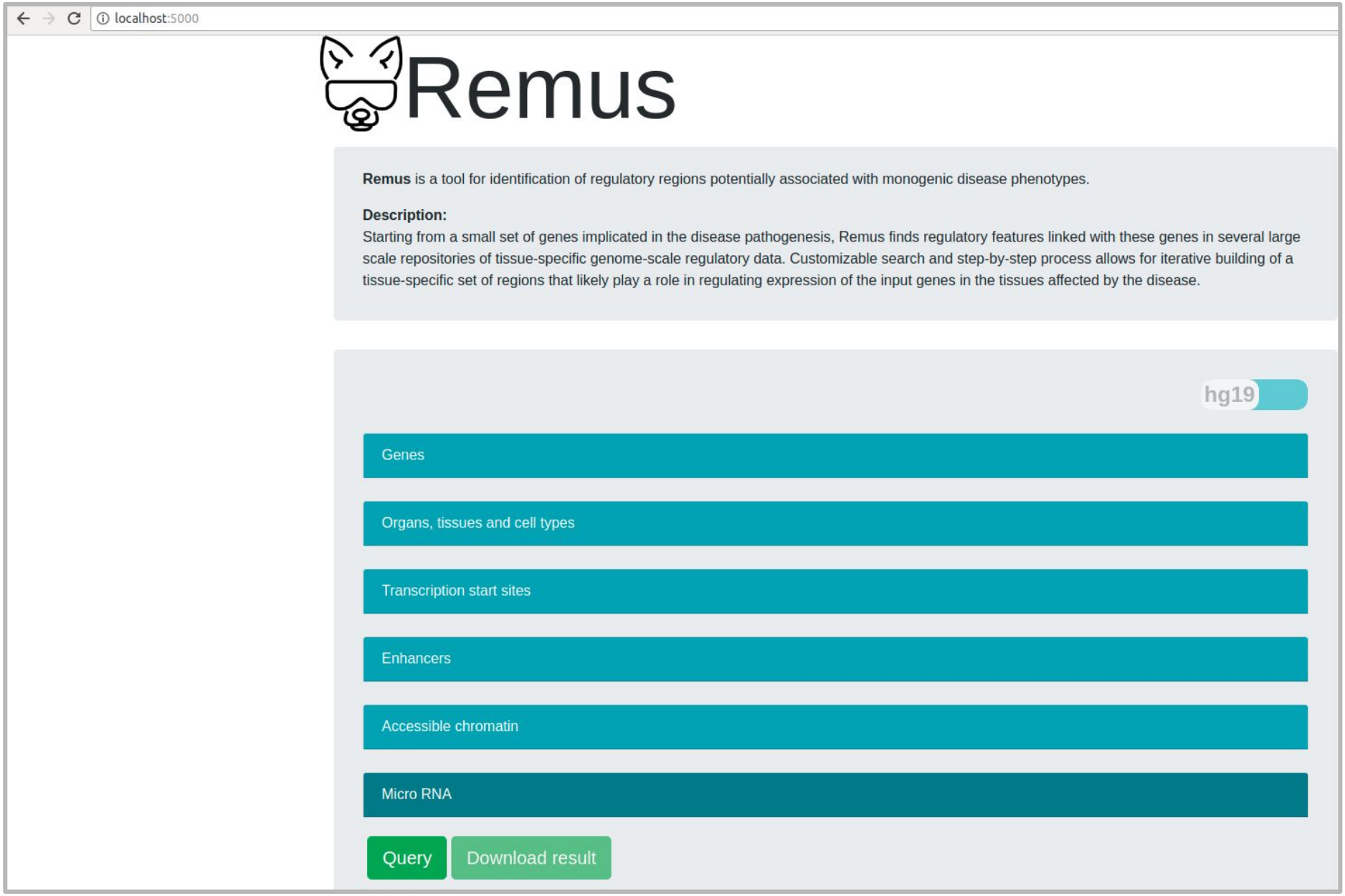
Remus - REgulatory MUtation Search

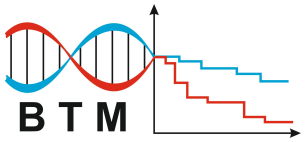

tissue-specific
Experimental, tissue-specific
DNase-seq (ENCODE)
&
RNA-seq (GTEx)
ChIP-seq
(ENCODE)
CAGE
(FANTOM5)
Regions of interest
Cis regulatory elements of
disease gene and its
regulatory neighborhood


BTM
ReMuS
MEDICAL UNIVERSITY OF LODZ
tissue-specific
Experimental & predicted
miRTarBase
DIANA-TarBase
RegNetwork
RegulatoryCircuits.org
genes regulating and regulated by disease gene(s)
Experimental, tissue-specific
DNase-seq (ENCODE)
&
RNA-seq (GTEx)
ChIP-seq
(ENCODE)
CAGE
(FANTOM5)
Tissue / cell-type specific filter
Regions of interest
Cis regulatory elements of disease gene and its regulatory neighborhood

BTM

MEDICAL UNIVERSITY OF LODZ





MEDICAL UNIVERSITY OF LODZ

hg19

Genes

Select genes:

×HNF1B HNF

HNF1A
HNF1A-AS1
HNF1B
HNF4A
HNF4G

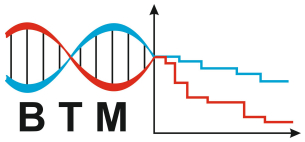
Transcription start sites

Enhancers

Accessible chromatin

Micro RNA

Query | Download result



MEDICAL UNIVERSITY OF LODZ

hg19

Genes

Organs, tissues and cell types

Select organs/tissues/cell types:

×kidney (CHRM, ENH_F5, TSS_F5)
×kidney epithelial cell (CHRM, ENH_F5, TSS_F5) kidne

kidney (CHRM, ENH_F5, TSS_F5)
kidney epithelial cell (CHRM, ENH_F5, TSS_F5)
kidney_embryonic (CHRM)
left kidney (CHRM)
left kidney_embryonic (CHRM)
right kidney (CHRM)

Enhancers

Accessible chromatin

Micro RNA

Query Download result

Transcription start sites

- [x] FANTOM5
  - ( ) Active in all selected tissues
  - (•) Active in any of selected tissues

bps upstream

3000

bps downstream

100

Enhancers

- [x] FANTOM5
  - ( ) Active in all selected tissues
  - (•) Active in any of selected tissues

Kb upstream

500

Kb downstream

500

- [x] ENCODE
  - ( ) Active in all selected tissues
  - (•) Active in any of selected tissues

Kb upstream

500

Kb downstream

500

Accessible chromatin

- [x] ENCODE
  - ( ) Accessible in all selected tissues
  - (•) Accessible in any of selected tissues

Kb upstream

500

Kb downstream

500

Enhancers

- [x] FANTOM5
  - ( ) Active in all selected tissues
  - (•) Active in any of selected tissues
  - Kb upstream: 500
  - Kb downstream: 500
- [x] ENCODE
  - ( ) Active in all selected tissues
  - (•) Active in any of selected tissues
  - Kb upstream: 500
  - Kb downstream: 500

Accessible chromatin

- [ ] ENCODE

Micro RNA

- [x] miRTarBase
  - [ ] Include weak interactions
- [ ] miRWalk
- ( ) Active in all selected tissues
- (•) Active in any of selected tissues

Query | Download result

Summary table

Time elapsed (s)	3.332624
No. features	7
No. base pairs	60226

Acknowledgments

Damian Skrzypczak

Wojciech Fendler

https://github.com/seru71/Remus

Project funded by National Science Center in Poland
(POLONEZ 2016/23/P/NZ2/04251)

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 665778.

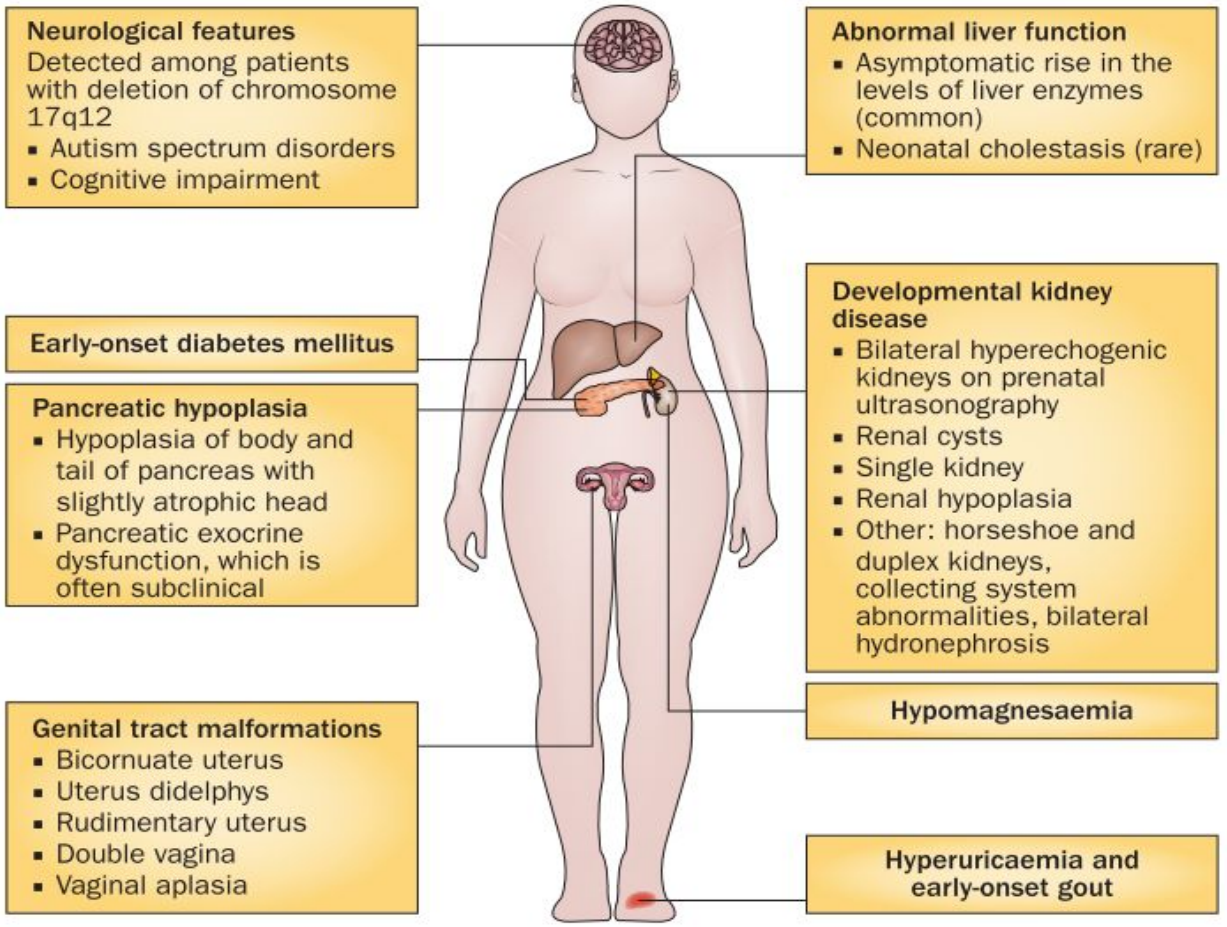
Study case: RCAD syndrome

Prevalence:

1-9 / 1 000 000

Autosomal dominant

HNF1B



Clissold, RL et al. Nature Reviews, 2014

Study case: RCAD syndrome

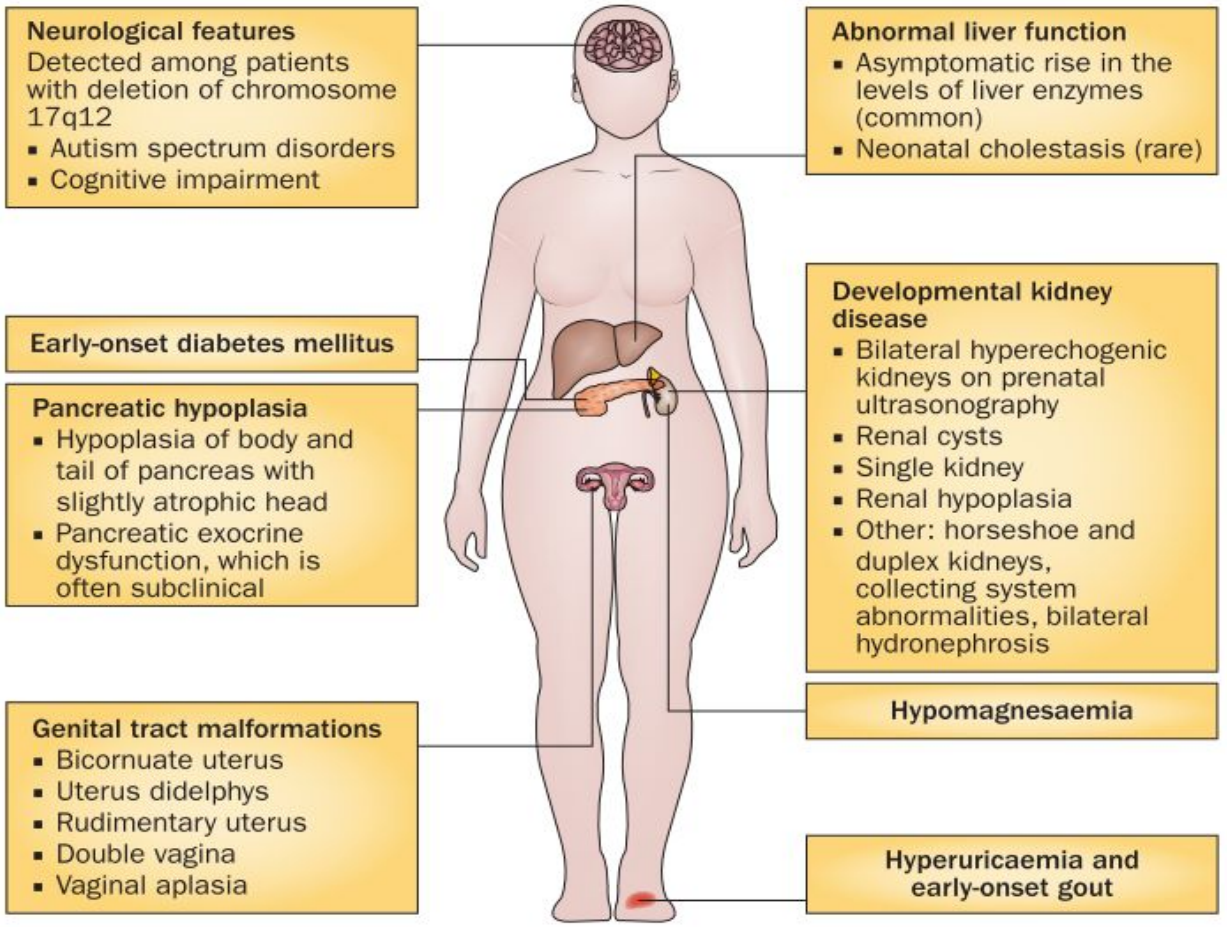
Prevalence:

1-9 / 1 000 000

Autosomal dominant

HNF1B

What if no mutations in HNF1B ?



Clissold, RL et al. Nature Reviews, 2014

Case: RCAD

WGS data for 9 patients with RCAD phenotype and no HNF1B mutatation

15+ additional WGS samples

Case: RCAD

WGS data for 9 patients with RCAD phenotype and no HNF1B mutatation

15+ additional WGS samples

Looking for interesting cases!

Acknowledgements

Arkadiusz Michalak

Beata Małachowska

Anna Hogendorf

Wojciech Fendler

Karolina Antosik

Paulina Młudzik

Maciej Borowiec

https://biostat.umed.pl/polonez_eng.html

Project funded by National Science Center in Poland
(OPUS 2014/15/B/NZ5/00144, POLONEZ 2016/23/P/NZ2/04251)

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 665778.

Integrative approach