

#### The Importance of Bioinformatics in the Detection of Cognitive Disorders

#### **Topics in Bioinformatics**



G Antwerp			Search	Q	I Nederlands	
Cognitive Genetics Centre of Medical Genetics University of Antwerp	5					
Research Interests	People	For Students	Voor de j	patiënt	Contact	
You are here: <u>UAntwerp</u> > <u>Research groups</u> > Co	gnitive Genetics			T 😢	weet F Like	47777



#### **Research Interests**

- > Fragile X syndome
- > Dynamic mutations
- > Novel syndromes and disorders
- > Gene interaction networks
- > Bioinformatics

#### Voor de patiënt

- > Patienten organisaties
- > Behandeling van het fragiele X syndroom

#### **Cognitive Genetics**

The term cognitive genetics was introduced to bridge the concepts of genetics and cognitive processes in the first decade of the 21st century. This relatively novel branch of genetics studies the influence of genetic variation on cognition and central nervous system disorders.

#### Mission

Our mission is to identify genetic causes of cognitive disorders and to study the molecular defects in order to eventually develop rational the rapy.

#### Embedding

Our research group Cognitive Genetics is part of the research cluster Medical Genetics of the Faculty of Pharmaceutical, Biomedical and Veterinary Sciences of the University of Antwerp. The research cluster, in combination with diagnostic and clinical units forms the Center of Medical Genetics. We are part of the University of Antwerp research excellence center GENOMED.



Maintained by the group:

#### VariantDB NGS data analysis suite

Galaxy@BioMina Customised microarray and NGS data analysis tools

#### CNV-WebStore MicroArray data analysis and interpretation

ADNP gene Phenotype information







#### **Cognitive Genetics**

The term cognitive genetics was introduced to bridge the concepts of genetics and cognitive processes in the first decade of the 21st century. This relatively novel branch of genetics studies the influence of genetic variation on cognition and central nervous system disorders.

#### Mission

Our mission is to identify genetic causes of cognitive disorders and to study the molecular defects in order to eventually develop rational therapy.

#### Databases

Maintained by the group:

VariantDB NGS data analysis suite

#### Galaxy@BioMina Customised microarray and NGS

data analysis tools

#### **CNV-WebStore**

## 6

#### Intellectual disability

#### Formerly "mental retardation" IQ below 70

- Diagnosed in childhood
- In combination with adaptive limitations











Universiteit Antwerpen

## Frequency of Intellectual disability



### Causes of Intellectual disability a reflection











# FISH

#### **Prader-Willi Syndrome**



### Labeling with fluorescent dye Denature & Hybridize Chromosoma

Fluorescence In Situ Hybridization

D	i	а	1	0

F3 Hier echt recentere plaatjes. Ook prader willi Frank; 29-10-2014



6

#### **Array-CGH**



Dia 1	1
-------	---

F4	Nieuwe dia
	Frank; 29-10-2014

## 6



**Array-CGH** 

## 6

ĵ,



• Two information channels: intensity and genotype



## 6

ĵ,



• Two information channels: intensity and genotype



## 6

ĵ,



• Two information channels: intensity and genotype















#### Different algorithms for Deletion/duplication testing

#### Table 1 Proportion of CNVs called by Pinto et al. detected by the different methods used

			Pi	nto et al.		
	De (n	letions = 378)	Du (r	plications n = 130)	(n	Overall = 508)
Algorithm	%	P-Val	%	P-Val	%	P-Val
2/3 Vote	81.5		69.6		74.4	
QuantiSNP	79.0	0.19	66.0	0.26	71.3	0.13
PennCNV	80.0	0.3	75.2	0.16	77.2	0.15
VanillalCE	83.0	0.31	55.1	0.0073	66.3	0.0023

Results are averaged over 181 HapMap samples. P-values are calculated using a one-tailed z-test for difference between proportions, comparing the majority vote against each separate method.

# But what if we find a deletion/duplication?

MOVE LEFT	MOVE RIGHT	ZOOM IN	ZOOM OUT
15% 50% 95%	15% 50% 95%	1.5x 5x 10x	1.5x 5x 10x
16p13.3 16p12.3	16p11.2 16q11.2 1	6q12.1 16q12.2 16q21	16q23.1



#### 2004: CNV in the control population



number variant

Janet Young,<sup>2</sup> Pär Lundin,<sup>3</sup> Susanne Månér,<sup>3</sup> Hillary Massa,<sup>2</sup> Megan Walker,<sup>2</sup> Maoyen Chi,<sup>1</sup> Nicholas Navin,<sup>1</sup> Robert Lucito,<sup>1</sup> John Healy,<sup>1</sup> James Hicks,<sup>1</sup> Kenny Ye,<sup>4</sup> Andrew Reiner,<sup>1</sup> T. Conrad Gilliam,<sup>5</sup> Barbara Trask,<sup>2</sup> Nick Patterson,<sup>6</sup> Anders Zetterberg,<sup>3</sup> Michael Wigler<sup>1</sup>\*

## CNV are stored in database of genomic variants

	Datal A curated ca	base of talogue of l	Genon	nic Va	riants ural variation
	About the Project Genome Browser	Downloads Query Tool	Links Submissions	Statistics Contact Us	FAQ Training Resources
Keyword,	Landmark or Example	Region Sos: RP11-34P13;	earch: CFTR, 7q11.21; c	:hr7:71890181-	Search GRCh37/hg19 ▼ •72690180
		Find DG by Study by Method by Platform	<b>W Variants</b> by Sample by Variant by Chromosom	e	
		Summa	ry Statistic	s	

 Stat
 Merged-level Sample-level

 CNVs:
 552586
 6359956

 Inversions:
 3164
 30446

 Number of Studies:
 72

News: May 2016 Update and Newsletter has been issued



#### Are there genes in the region?

#### Genome is annotated in UCSC



## Are there relevant genes in the region? Disease genes are annotated in the OMIM

	entry/610112	☆ 🖪
👖 Apps 🖉 CNV-WebStore 🛞 Vari	iantDB 🍺 DECIPHER 💈 OMIM 🔗 PM 🔫 Galaxy 🗋 BioMina LTS 🛄 Trello 🕒 UA 🕁 Cognitive Genetics 🍥 Prezi 💪 UA CP 🝐 Google	Drive G Google »   📙 Andere bladwijzer
Home About Statistics • Down	loads • Help • External Links Terms of Use • Contact Us MIMmatch Donate •	
Search OMIM	Search	
Advanced Search *		
Table of Contents for *610112		External Links
Title	*610112	<ul> <li>Genome</li> </ul>
Text	C-MAF-INDUCING PROTEIN	► DNA
Gene Function		► Protein
Gene Structure	Alternative titles: symbols	Gene Info
Molecular Genetics	CMIP	<ul> <li>Clinical Resources</li> </ul>
References	KIAA1694	<ul> <li>Variation</li> </ul>
Edit History		<ul> <li>Animal Models</li> </ul>
MIMmatch (login)	Other entities represented in this entry:	<ul> <li>Cellular Pathways</li> </ul>
	C-MAF-INDUCING PROTEIN, TRUNCATED, INCLUDED; TCMIP, INCLUDED	
	HGNC Approved Gene Symbol: CMIP	
	Cytogenetic location: 16q23.2-q23.3 Genomic coordinates (GRCh38): 16:81,445,169-81,711,761 (tom NCB)	
	TEXT	
	Cloning and Expression	
	By screening for genes with the potential to encode large proteins expressed in brain, Nagase et al. (2000) cloned	
	KIAA1694. The predicted protein contains 757 amino acids. RT-PCR ELISA detected ubiquitous expression of	
	KIAA1694, with highest levels in brain, followed by ovary and kidney. Expression was high in all brain regions	

## Do the parents have the same deletion?

#### Father

#### Mother



## Did we find the same deletion before? Build and search in house database



## Did others find the same deletion before? Search DECIPHER

Location	▲ Gene	Allele	Transcript	0 Co	nsequence	⊜ Inh Ge	eritance notype	Pathogenicity     Contribution ?	Sh	ared ?	Links	E.
No data availa	ble in table											
										Add Sequer	nce Var	riant
Browser	Genes 3	Overlappin	g Patients 26	Overlapping	g Syndromes 🟮							
≡ Tracks	Chr 16	p13.3	pt	2.3 p12.2 p12.1	p11.2	q11.2	q12.1 q12.2	q <b>21</b> q22.1	q23.1	q24.1		< >
Genes	0 Mb	81.35 Mb GAN >	81.40 Mb	81.45	S Mb CMIP >	1.50 Mb	81.55 Mb	81.60 Mb	81.65 Mb	81.70 Mb	×	<b>Q Q</b>
and the fill												+ [] • •
Genes Legend	% HI range	es:	20 30	40 50	60 70	80	90 100	No % HI score				t =
Morbid Genes	_	GAN >		•							«	Э
dadada												9
Sequence	>>										*	2
This Patient: Copy-Number Variants						Affected patien					*	6
DECIPHER: Copy-Number										2	978 **	
Variants			1724			3	06151				_	
1000			4729	289101		33102						
											0000000	

#### **CNV-WebStore a graphical interface**

	TeraStation 난 UA 🗶 Cobb 📩 Met	ropolis 🗋 VBF 🔄 Antwerpen file 🕨 NP	RC 🚾 AskOxford 📄 Rongga's 📐 Ne	spresso 🗋 Weberclub 🛞 Jiba 📄	Prisma 🔧 Goo
Store: Online CNV A	nalysis, ÷				
CNV-	WebStore	An Online CNV	/ Analysis, Storage and Interpreta	tion Platform Centre of M	V 💙 🔪 Medical Gene
Main   New	Analysis   Browse Results   Expo	rt   Settings   Other Tools   Documen	itation   Contact   Print		Frank  🖷
Browse Pr	ojects   Search : Sample/Region/Go	ane/Phenotype   Filter   Set Permissi	ons	Using b	uild NCBI 36 (hg1
Search	detected aberrations	based on chromosomal	region / Gene Symbo	1	
Enter the ch	romosomal region in the field below	w en click 'search'. The region should	be specified in UCSC-style, by ch	romosome band or by RefSeq	gene symbol, e
chr7:7500000	-9200000, 4q21.3, 16p11.2-p12.2 or S(	DX12.			
Search region	: 7q11.23 Visualise	by 🔍 image or 🔘 table. search			
	s Containing CNV's in	7011.23			
Samples	sponds to chr7:71,800,000-77,400,000	/4=3			
Sample: Query corres					
Sample: Query corres	MOVE LEFT	MOVE RIGHT	ZOOM IN	ZOOM OUT	
Sample: Query corret	MOVE LEFT	MOVE RIGHT 15% 50% 95%	ZOOM IN 1.5x 5x 10x	200M OUT	
Sample: Query corre.	MOVE LEFT	MOVE RIGHT	ZOOM IN 1.5x 5x 10x 7422-11 243344	ZOOM OUT	
Sample: Query corre 258.0 Kb	MOVE LEFT	MOVE RIGHT	ZOOM IN 1.5x 5x 10x 2422-11 7401-1	ZOOM OUT 1.5x 5x 10x 1.5x 744 ZS3	Preferenc

Vandeweyer et al. BMC Bioinformatics 2011, 12:4 http://www.biomedcentral.com/1471-2105/12/4



SOFTWARE

Highly accessed Open Access

### CNV-WebStore: Online CNV Analysis, Storage and Interpretation

Geert Vandeweyer<sup>1</sup>, Edwin Reyniers<sup>1,2</sup>, Wim Wuyts<sup>2</sup>, Liesbeth Rooms<sup>1</sup>, R Frank Kooy<sup>1\*</sup>

#### Abstract

**Background:** Microarray technology allows the analysis of genomic aberrations at an ever increasing resolution, making functional interpretation of these vast amounts of data the main bottleneck in routine implementation of high resolution array platforms, and emphasising the need for a centralised and easy to use CNV data management and interpretation system.

**Results:** We present CNV-WebStore, an online platform to streamline the processing and downstream interpretation of microarray data in a clinical context, tailored towards but not limited to the Illumina BeadArray platform. Provided analysis tools include CNV analsyis, parent of origin and uniparental disomy detection. Interpretation tools include data visualisation, gene prioritisation, automated PubMed searching, linking data to several genome browsers and annotation of CNVs based on several public databases. Finally a module is provided for uniform reporting of results.

**Conclusion:** CNV-WebStore is able to present copy number data in an intuitive way to both lab technicians and clinicians, making it a useful tool in daily clinical practice.

## Analysis using CNV-WebStore









European Journal of Medical Genetics 52 (2009) 94-100





#### Contents lists available at ScienceDirect European Journal of Medical Genetics

2 EUROPEAN JOURNAL

journal homepage: http://www.elsevier.com/locate/ejmg

Original article

Fourteen new cases contribute to the characterization of the 7q11.23 microduplication syndrome

Nathalie Van der Aa<sup>a,\*</sup>, Liesbeth Rooms<sup>a</sup>, Geert Vandeweyer<sup>a</sup>, Jenneke van den Ende<sup>a</sup>, Edwin Reyniers<sup>a</sup>, Marco Fichera<sup>b</sup>, Corrado Romano<sup>b</sup>, Barbara Delle Chiaie<sup>c</sup>, Geert Mortier<sup>c</sup>, Björn Menten<sup>c</sup>, Anne Destrée<sup>d</sup>, Isabelle Maystadt<sup>d</sup>, Katrin Männik<sup>e</sup>, Ants Kurg<sup>e</sup>, Tiia Reimand<sup>f</sup>, Dom McMullan<sup>g</sup>, Christine Oley<sup>g</sup>, Louise Brueton<sup>g</sup>, Ernie M.H.F. Bongers<sup>h</sup>, Bregje W.M. van Bon<sup>h</sup>, Rolph Pfund<sup>h</sup>, Sebastien Jacquemont<sup>i</sup>, Alessandra Ferrarini<sup>i</sup>, Danielle Martinet<sup>i</sup>, Connie Schrander-Stumpel<sup>j</sup>, Alexander P.A. Stegmann<sup>j</sup>, Suzanna G.M. Frints<sup>j</sup>, Bert B.A. de Vries<sup>h</sup>, Berten Ceulemans<sup>k</sup>, R. Frank Kooy<sup>a</sup>



#### Williams syndrome: a common deletion in the 7q11.23 region

Deletion of 1.55 Mb contiguous gene syndrome



## Analysis using CNV webstore

#### Chromosome 7





## Williams-Beuren syndrome: a common deletion in the 7q11.23 region





REPORT

## The Contribution of *CLIP2* Haploinsufficiency to the Clinical Manifestations of the Williams-Beuren Syndrome

Geert Vandeweyer,<sup>1</sup> Nathalie Van der Aa,<sup>1</sup> Edwin Reyniers,<sup>1</sup> and R. Frank Kooy<sup>1,\*</sup>

	10	20	30	40	50	60	70	80	90	100	110	120
Proximal AluSz	GGCTGGGTAT	GGTGGCTCATG	CCTGTAATCCC.	AGCTCTTTGGG	AGGCCGAGGA	GGGCAGATCACT	TGAGGTCAG	GAGTTCGAGA	CCAGCCTGGC	CAACATGGTO	GAAACCTCGTC	TCTACT
Distal AluSz	GGCCAGGCGT	GGTGGTTCACA	CCTGTAATCCC.	AGCACTTTGGG	AGGCTGAGGC	AGGCAAATCACT	TGAGGTCAG	GAGTTCGAGA	CCAGCCTGGC	CAACATGGTO	GAAACCCCATC	TCTACT
	10	20	30	40	50	60	70	80	90	100	110	120
	130	140	150	160	170	0	190	200	210	220	230	
Proximal AluSz	AAAAATAACAA	A-AATTAGCCO	GGCGCGACGGC	ACATGCCTGT	AATCCCAGCT	ACTCAGG AGGCT	GAGGCAGGAG	GAATGGCTTG.	A. CC FGGAAG	GCAGAGGTTG	TAGTGAGCCA	AGTTTA
Distal AluSz	AAAAGTACAAA	ACAATTAGCTO	GGTTGGT-GGT	GTGGGTCTGT	AATCCCAGCT	GTTCAGGGAGCT	GAGGCAGGAG	GAATGGCTTG.	AJ COCGGGAG	GCAGAGGTGA	CAGGGAGCCA	AGATCA
	130	140	150	160	170	0	190	200	210	220	230	
24	0 250	2.60	270	280	290							
Proximal AluSz	CACCATTGCAC	TCCAGCCTGGG	CAACAAAGGAA	GACTCTGTCTC	CATAAAAATA	AAGA						
Distal AluSz	CGCCACTGCAC	TCCAGCCTGGG	CAACAGAGCGA	GACTCCGTCTC	AAAAAAAAA	AAAA						
24	0 250	260	270	280	290							

### Five patients with a chromosome 1q21.1 triplication show macrocephaly, increased weight and facial similarities



Anke Van Dijck <sup>a, b, \*</sup>, Ilse M. van der Werf <sup>a</sup>, Edwin Reyniers <sup>a, b</sup>, Stefaan Scheers <sup>a, b</sup>, Meron Azage <sup>c</sup>, Kiana Siefkas <sup>d</sup>, Nathalie Van der Aa <sup>a</sup>, Amy Lacroix <sup>e</sup>, Jill Rosenfeld <sup>f</sup>, Bob Argiropoulos <sup>g, h</sup>, Kellie Davis <sup>g</sup>, A.Micheil Innes <sup>g, h</sup>, Heather C. Mefford <sup>d, e</sup>, Geert Mortier <sup>a, b</sup>, Marije Meuwissen <sup>b</sup>, R.Frank Kooy <sup>a</sup>

European Journal of Medical Genetics 58 (2015) 503-508



### CNV is found in 10% of all patients with Intellectual disability



Figure 4. Overview of all CNVs reported in genome-wide microarrays studies in mental retardation. CNVs are represented by colored bars. Copy number losses and copy number gains are depicted, respectively, on the left-hand side and the right-hand side of the chromosomes. 1) Red bars, CNVs associated with a well-known OMIM syndrome; 2) orange bars, novel recurrent CNVs; 3) green bars, de novo CNVs not known to the previous categories; and 4) blue bars, CNVs of unknown clinical significance.

### Causes of Intellectual disability a reflection



### Draft sequence of the Human Genome Human Genome Project (1990 – 2001)

#### Public consortium



Celera Private company



February 2001



## 6

## NGS Data analysis

#### What kind of data are we working with?

#### From sequence to variant: Analysis flow

From variant to knowledge: Interpretation flow





#### What kind of data are we working with?

- Sanger Sequencing:
  - 1 amplicon / reaction
  - 1 sequence / amplicon (or 2)
  - Visual inspection for overlapping peaks



Father: heterozygous carrier, c251C>T, codon ACC>ATC, p.T84I



#### What kind of data are we working with?

- Sanger Sequencing:
  - 1 amplicon / reaction
  - 1 sequence / amplicon (or 2)
  - Visual inspection for overlapping peaks



- Next-Generation Sequencing:

- Massive Parallel sequencing
- exome panel: > 200.000 targets (all exons in the genome)
- Multiple amplicons / target
- optimal design: > 40 unique
  - fragments covering *every* nucleotide in targets.



#### What kind of data are we working with?



=> Amount of data : > 8.000.000 sequences / sample=> Impossible to manually inspect all reads for mutations.

6

## NGS Data analysis

#### What kind of data are we working with?

- Data format : FASTO

- FASTA :

>Sequence\_Name AACTACTAGATACTGATAGTATATCTCTCTTAATCGA GCTCTAGATCGATCTATACCGAT

- Add Quality (fasta-Q => FASTQ) @Read\_Name AACTACTAGATACTGATAGTATATCTCTCTTAATCGA + BCEECEEFFECGECGECFGFF@?<<=??<>53@##



#### What kind of data are we working with?

- Data format : FASTQ @Read Name AACTACTAGATACTGATAGTATATCTCTCTTAATCGA + BCFFCFFFFCGFCGFCFGFF@?<<=??<>53@##





#### From sequence to variant: Analysis flow



#### From sequence to variant: Analysis flow

![](_page_50_Figure_2.jpeg)

Low quality leads to high error rates (cfr Phred Score) => We want a limit of 1 error in 1000 positions => Due to chemical degradation, 3' ends are lower quality => Trim everything on 3' end with quality < 30

![](_page_51_Picture_0.jpeg)

#### From sequence to variant: Analysis flow

Sequence – To – Variant **Read Mapping** 

![](_page_51_Figure_5.jpeg)

6

### NGS Data analysis

#### From sequence to variant: Analysis flow

Sequence – To – Optimize Mapping Variant

Remove Duplicate reads (picard)
 Reduce computational time
 Reduce amplification bias

- Realign around indels (GATK)

![](_page_52_Picture_6.jpeg)

6

### NGS Data analysis

#### From sequence to variant: Analysis flow

Sequence – To – Call And Annotate Variants Variant

- <u>Call Variants (GATK)</u>

- Search for positions with statistically significant evidence for a non-reference nucleotide

- Take into account: base-quality, position in read, strand bias, ...

![](_page_53_Picture_7.jpeg)

![](_page_54_Picture_0.jpeg)

#### <u>Galaxy</u>

- A website offering an easy way to run complete pipelines
- No programming skills needed, very usefull for dynamic analysis

![](_page_54_Figure_5.jpeg)

## 6

## **Genetic Variation**

- Single Nucleotide Polymorphism
  - Fixed position in the genome that \*might\* differ in sequence between individuals
  - 1 SNP / 700bp
  - Long thought to be major source of variation

![](_page_55_Figure_6.jpeg)

6

## NGS Data analysis

#### From sequence to variant: Analysis flow

Sequence – To – Call And Annotate Variants Variant

Annotate Variants (ANNOVAR, snpEff, ...) - Add as much information to the variant as possible to ease interpretation - Effect on Gene transcription (RefSeq, Ensembl, UCSC) - Quality parameters (GATK) - Occurence in control populations (dbSNP, ESP, HapMap, 1KG, ...) - Known pathogenic variations (dbSNP, OMIM, ...) - Effect on gene function (PolyPhen, MutationTaster, Sift, ...)

![](_page_57_Picture_0.jpeg)

#### From variant to knowledge: Interpretation flow

Effect on gene function (~ from high to low)

- Variant causes gain/loss of stop/start coding?
- Variant causes aberrant splicing of the transcript?
- Variant replaces a highly conserved nucleotide/amino acid ?
- Variant replaces an aminoacid, and is not reported in control populations ?
  - Variant can modify binding of regulatory elements?

 $\Rightarrow$  Extended annotation is critical

 $\Rightarrow$  Manual inspection of > 20.000 variants/sample is impossible.  $\Rightarrow$  Again: automation is needed

#### **VariantDB** Web-based interface

				Type the sample name	F7_Index			Download : bam bai, vcf.gz, vcf.gz lbi
				Filter Settings Annotation	ions Export Statistical Char	rts Sample Log		
				Select Annotation	is To show			Preset Annotation
				Hover the mouse over available a	annotations to get more information.			ALL_GATK #
				GATK_Annotations Information	lon			Effect On RefSeq g
/CF to VariantDB (version 0.1.2)				AltelicRatio	Alt_Allele_Depth	Base_Quality_Rank_Sum	Fisher_Strand_Blas	SnpEff_Grch37.66 m
				Genotype	Mapping_Quality	Mapping_Quality_Rank_Sum	Phred_Genotype	🗐 all g
ACF THE:				Phred_Polymorphism	Cuality_By_Depth	C Read_Position_Rank_Sum	C Ref_Allele_Depth	GO_anno 🕿
5: Unified Genotyper on data 68 and data 30 (VCF) 🔹				ANNOVAR Information	Contrancies_riter			U WebTools @
nified Genotyper VCF File				1000g2012apr_afr	🗏 1000g2012apr_all	1000g2012apr_amr	🗎 1000g2012apr_asn	Add Annotations
store VCF and BAM Files::				1000g2012apr_eur	CADD_phred	CADD_raw	Ensembl_Exon	Save Current Annotation
Mark T				Ensembl_GenelD	Ensembl_GeneLocation	Consembl_TranscriptiD	Casemia SuperDura Seere	Save Current Apostations
ies '	a	- Annes		RefSeg Exon	RefSeg GeneLocation	RefSeg Protein Length Differer	ce RefSeg Symbol	Sare Guitem Amountons
his option allows you to send the BAM and VCF files to our storage server for dynamic loading into IGV. If you	store them there	please		RefSeq_Transcript	RefSeq_VariantType	RefSeq_cPointAA	RefSeg_cPointNT	
elete them here.				UCSC_Exon	UCSC_GeneLocation	UCSC_Symbol	UCSC_Transcript	
AM File.:				UCSC_VariantType	UCSC_cPointAA	UCSC_cPointNT	🖾 esp5400_aa	
4: Final BAM In Linkage				esp5400_all	🖾 esp5400_ea	esp6500_aa	esp6500_all	
T. FINAL DAWL AT LOTINAGE .				esp6500_ea	U IJb_GERP	Ub_LRT	Ib_MutTast	
Provide a Sample Name ::				C IID_PhyloP	IjD_PolyPhen2     sec136_MAE	in inp_sm	snp136_rsiD	
Type the samplename				snp137_Clinical	snp137_MAF	snp137_NrChr	snp137_rsiD	
Type the samplename				WebTools Information		- supres general		le se la companya de
no name is specified, a new sample will be created, and you will be notified of the name			Use This Sample/Region				Data Visualisation in IGV	
Sample Name.:			Type the sample name	F7_Index			Download : bam bai, ysf.gz, ysf.gz, tbi	
DemonstrationSample			Filter Settings Annotations	Export Statistical Charts	Sample Log			-
Samela Candan .			Build your query				Preset Filters	
ample Gender: :							Cuality #	
Female 🔹			Negate Filter On		Argument	Values	HaloPlex_Quality to	
his can be set from the database frontend as well.			Filter On Family Information		[C2:Calbar]		🔲 not in snp130 🗑	
ariantDR-Server '			- Not Match T		F7_Mother	As Any Ganchina	🖂 common filters 😭	
			(incasting)	2		Partit Centrality	Apply Filter	
Main Server @ University of Antwerp 🔻			Ether On Operations (ethernation		O			
pecify the VariantDB server you wish to send the data to. You MUST have a valid account on the target server	r, identical to your	account	- Match Y Rel Occ Con	trol Samples (Any Genotype)	Smaller Or Foual Than *	0.05	Sava Current Filter	
are.			Ellar De Laratias Johannatias		Contraction of the second second	10.00	Save Current Filter	
			Filter On Effect On Transcript Information				Save Current Filter	
Execute			Filter On Genotype Composition Information	tion 💡				
			Filter On Quality Information					
			- Match • Quality By De	pth 🔹	Bigger Or Equal Than 💌	4.8		
hat it does			Filter On Mutation_Effect_Predictions Info	rmation 💂				
			Filter On snpEff, Annotations, from, GRCI	37_66 Information				
is coordinate of the second seco	compared betwe	211	Filter On ClinVar Information 💡					
imples, filtered on various annotations etc. To add servers, specify them in the tool configuration XML file.			Filter On Gene_Ontology Information					
nut file			<u>.</u>					
put me	Location	Ref.Allele		Alt.Allele	Ref.Depth	Alt.Depth Allel	c Balance	Genotype
CF file from the GATK Unified Genotyper.	chr1:45973938	C		Т	18	13 0.41	24	Heterozygous
	Validity & Significance	Base_Quality_Rank_Sum: -0.5 Phred_Genotype: 99	Fish Phra	ier_Straid_Blas: 0 ed_Polymorphism: 344,19	Mapping_G Quality_By	Juality: 60 _Depth: 11.1	Mapping_Quality_Rank_Si Read_Position_Rank_Sum	um: 0.781 x: -1.021
	Set inheritance : Not Defined	Strand_Bias -185.25		City Comment				Long Harden Barbard
Jutputs	Validation -	everlap copy number p	gain Uncertain significan	ce classified by single su	ubmitter muttigte conditi	ions	MMACHC	2013-11-19

Use This Sample/Region

Text file with some results from the vcf-parser.

45973938	C			т		18		13	0.4194		Hetero	zygous	
ity & Significance	Base_Quality_Rank_Sun Phred_Genotype: 99 Strand_Blas: -185.25	£ -0.5	Fisher_ Phred_	"Strand_ Blas: 0 Polymorphism: 344,19			Mapping_Qs Quality_By_I	ulity: 60 Depth: 11.1		Maps Read	oing_Quality_Rank_Sum: 0.781 {_Position_Rank_Sum: -1.021		
/ info: - (0p/1sit)	CV_Match CV_	VarType	Class	Class_Co	mment		Disease			Gene	Gene_Effect	Last_Update	PubMed
ANC -	overlap coj	ry number gain	Uncertain significance	classified	by single submitter		multiple conditio	03		MMACHC		2013-11-19	
	overlap coj	iy number gain	Uncertain significance	classified	by single submitter		Autism			CCDC17		2013-11-19	
	exact sin	gle nucleotide variant	Pathogenic	classified	by single submitter		Methylmalonic a	cidemia with homocystinutia		MMACHC	STOP-GAIN	2014-03-30	PubMed
	Panel Definition			Panel_Gene				Panel Gene Comment					
	ID genes			MMACHC				OMIM 609831 AR -					
	RefSeq_Exon	RefSeg_GeneLocation	RefS og Protein	Length_Difference		RefSeq_Symi	h for	letSeg_Transcript	RefSeq_Variant	Type	RefSeq_cPointAA	RefSeq_cP	oin#VT
	3	exonic	0			MMACHE		NM_015506	stopgain SNV		p.R111X	e.C331T	
	Amino_Acid_Change_37	66 Codon_Change_37.66	Effect_37.66	Effect_Impact_37.66	Functional_Class	37.66	Gene_Coding_37.66	Gene_Exon_37.66	Ġe	ne_Symbol_37.66	Gene_Transcript_37.66	Transcript_B	oType_37.66
			DOWNSTREAM	MODIFIER			CODING		PI	RDX1	ENST00000262746	pietain_oodir	10
			DOWNSTREAM	MODIFIER			CODING		PI	RDX1	ENST00000319248	protein_codir	10
			DOWNSTREAM	MODIFIER			CODING		PI	RDX1	ENST00000372079	protein_codir	10
	R111*	Cga/Tga	STOP_GAINED	HIGH	NONSENSE		CODING	exon_1_45973084_45974	036 M	MACHO	ENST00000401061	protein_codir	10
			UPSTREAM	MODIFIER			CODING		м	MACHC	ENST00000474382	nonsense_mr	diated_decay
	×.		UPSTREAM	MODIFIER			CODING		M	MACHO	ENST00000477188	processed_tre	enscript

#### Universiteit Antwerpen

Data Visualisation in IGV

## 6

- Global Developmental Delay
- Failure to Thrive
- Autism Spectrum Disorder
- Cardiac: Atrial Septal Defect
- White Matter Lesions
- Facial Dysmorphism

## Young boy

![](_page_59_Picture_8.jpeg)

## Next Generation Sequencing Trio approach

![](_page_60_Figure_1.jpeg)

#### A de novo paradigm for mental retardation

Lisenka E L M Vissers<sup>1,2</sup>, Joep de Ligt<sup>1,2</sup>, Christian Gilissen<sup>1</sup>, Irene Janssen<sup>1</sup>, Marloes Steehouwer<sup>1</sup>, Petra de Vries<sup>1</sup>, Bart van Lier<sup>1</sup>, Peer Arts<sup>1</sup>, Nienke Wieskamp<sup>1</sup>, Marisol del Rosario<sup>1</sup>, Bregje W M van Bon<sup>1</sup>, Alexander Hoischen<sup>1</sup>, Bert B A de Vries<sup>1</sup>, Han G Brunner<sup>1,3</sup> & Joris A Veltman<sup>1,3</sup>

Nature Genetics 42, 1109-1112 (2010)

![](_page_61_Figure_0.jpeg)

6

## NGS analysis

Human hg 19	-	r20 👻 d	hr20:49,508,733-49,508,773	Go 👚 ·
	AME AIATYPE AIAFILE	p12.3 p12.1 p11.7	22 q11.1 q11.23 q13.12 42 bp	<b>q13.2 q13.</b> 49.508.770 bp
F9_Index_11121294.Variants	200			
HISEQ				
F9_Father.Readather.Variants			No Variants Found	
F9_Mother.Reather.Variants			No Vaviante Found	
HISEQ			No variants round	
F9_Index_111294.Reads,F9_Ind 11294.Variants Coverage		0-133]		
F0_Index_111204.Reads.F0_Ind 11204.Variants				
		0-96]		
r 9_r ather. Reads, r9_r ather. Van ts Coverage				
F9_Father.Reads.F9_Father.Vari ts				
F9_Mother.Reads,F9_Mother.Va ts Coverage		0- 112]		1981 Juni 2001 Juni 2002 Juni 2003 Juni 2
F9_Mother.Reads,F9_Mother.Va ts		<hr/>		
Sequence 🗕		ATCCATCTCATGCT I H L M L S I S C N P S H A		
RefSeq Genes			ADNP	K P N
11 tracks loaded	labe 20	508 773		263M of 409M

### NGS analysis: a single *de novo* mutation

![](_page_63_Figure_1.jpeg)

- 4bp del in ADNP
- Causes frameshift introducing stop codon
- Mutation not control databases, e.g.,ESP, 1000g, dbSNP

#### Universiteit Antwerpen

## Activity-Dependent Neuroprotective Protein

- Expressed in brain
- Zinc fingers/homeobox domain: potential transcription factor
- Involved in neurogenesis
- Involved in heart development
- Homozygous KO mice are embryonically lethal
- Heterozygous KO mice have cognitive & behavioral problems

# Ten patients with truncating mutations in *ADNP*

			Screening		Cohort	mutation in genomic	mutation in cDNA		Mutation	
Patient	Patient ID	Origin	method	<b>Cohort composition</b>	size	DNA (chr20)	(NM_015339.2)	Protein	Туре	Inheritance
1	111294	Antwerp	WES	Moderate to severe	10	g.49508752_49508755	c.2496_2499delTAAA	p.Asp832Lysfs*80	Frameshift	de novo
				ID and/or autism + dysmorphic features		delTTTA				
2	11-08612	Nijmegen	WES	Non-syndromic severe ID	100	g.49510040G>T	c.1211C>A	p.Ser404*	Nonsense	de novo
3	12130.p1	Seattle	WES <sup>2,16</sup>	ASD from the Simon Simplex Collection	189	g.49510028_49510029 delTT	c.1222_1223delAA	p.Lys408Valfs*31	Frameshift	de novo
4	1050237	Westmead	WES	Non-syndromic severe ID	95	g.49509086_49509098 delATTTGCTCGTAAG	c.2153_2165delCTTAC GAGCAAAT	p.Thr718Glyfs*12	Frameshift	de novo
5	3061-08D	Stockholm	WES	Moderate to severe ID and/or autism + dysmorphic features	45	g.49509094G>C	c.2157C>G	p.Tyr719*	Nonsense	de novo
6	122793	Antwerp	HRM	Autism	148	g.49508757_49508760 delTTAA	c.2491_2494delTTAA	p.Lys831llefs*81	Frameshift	de novo
7	07-06960	Nijmegen	MIPS	ID and/or autism	2743	g.49508443delG	c.2808delC	P.Tyr936*	Frameshift	de novo
8	2376	Troina	MIPS	ID and/or autism	Idem as patient 7	g.49508757_49508760 delTTAA	c.2491_2494delTTAA	p.Lys831llefs*81	Frameshift	de novo
9	2533	Troina	MIPS	ID and/or autism	Idem as patient 7	g.49509321G>A	c.1930C>T	p.644Arg*	Nonsense	parents not available
10	13545.p1	Seattle	MIPS <sup>16</sup>	ASD from the Simon Simplex Collection	2446	g.49509094_49509095 insT	c.2156_2157insA	p.Tyr719*	Frameshift	de novo

## **Dysmorphic features**

Phenotype	Frequency
Prominent forehead	5/8
High hairline	7/8
Eversion/notch eyelid	3/7
Hypertelorism	1/8
Broad nasal bridge	6/8
Short nose	2/8
Thin upper lip	6/7

![](_page_66_Picture_2.jpeg)

![](_page_66_Picture_4.jpeg)

### Clinical characteri

Phenotype	Frequency
Autism Spectrum Disorder (ASD)	10/10
Intellectual Disability (ID)	10/10
Developmental delay (motor)	9/10
Developmental delay (speech)	8/9
ADHD	2/9
Hypotonia	7/9
Growth retardation / Short stature	5/8
Feeding problems	5/8
Recurrent infections	5/8
Congenital heart defect	3/8
Hyperlaxity	6/8
Obesity	4/7
Hypermetropia	6/6
Seizures	2/7
Behavior	5/7
Insensitivity to pain	2/5
MRI brain abnormality	5/9
Hand abnormalities	6/8
Constipation	2/6

## 6

## **Statistics**

- The frequency of truncating *de novo* mutations in *ADNP* is significantly higher (p: 0.001852, odds ratio 13.24668, one-sided Fisher's exact test) in patients compared to the ESP cohort and additional controls from the Simons Siblings.
- The probability of detecting 8 or more *de novo* truncating events in *ADNP* within our cohort by chance was estimated to be p = 2.65e<sup>-18</sup> (binomial test) under a *de novo* rate of 1.2 non-synonymous coding variants per individual according to a probabilistic model of a locus specific enrichment for truncating variation [O'Roak et al., 2012].

### De novo mutations in ADNP cause new autism syndrome

Autism spectrum disorder Intellectual Disability Facial dysmorphology Congenital malformations

may explain etiology of 0.17% of ASD patients (95% confidence interval: 0.083–0.32%)

#### genetics

#### A SWI/SNF-related autism syndrome caused by *de novo* mutations in *ADNP*

Céline Helsmoortel<sup>1</sup>, Anneke T Vulto-van Silfhout<sup>2,16</sup>, Bradley P Coe<sup>3,4,16</sup>, Geert Vandeweyer<sup>1,5</sup>, Liesbeth Rooms<sup>1</sup>, Jenneke van den Ende<sup>6</sup>, Janneke H M Schuurs-Hoeijmakers<sup>2</sup>, Carlo L Marcelis<sup>2</sup>, Marjolein H Willemsen<sup>2</sup>, Lisenka E L M Vissers<sup>2</sup>, Helger G Yntema<sup>2</sup>, Madhura Bakshi<sup>7</sup>, Meredith Wilson<sup>8</sup>, Kali T Witherspoon<sup>3,4</sup>, Helena Malmgren<sup>9</sup>, Ann Nordgren<sup>9</sup>, Göran Annerén<sup>10</sup>, Marco Fichera<sup>11,12</sup>, Paolo Bosco<sup>13</sup>, Corrado Romano<sup>14</sup>, Bert B A de Vries<sup>2,15</sup>, Tjitske Kleefstra<sup>2,15</sup>, R Frank Kooy<sup>1</sup>, Evan E Eichler<sup>3,4</sup> & Nathalie Van der Aa<sup>1,6</sup>

#### Universiteit Antwerpen

Helsmoortel et al., Nat Genet 2014

#### LETTERS

Céline Helsmoortel Ilse van der Werf Anke Van Dijck

### Acknowledgements

#### **Geert Vandeweyer**

Nathalie Van der Aa

Liesbeth Rooms Jenneke van den Ende

![](_page_70_Picture_5.jpeg)

![](_page_70_Picture_6.jpeg)

![](_page_70_Picture_7.jpeg)

![](_page_70_Picture_8.jpeg)

![](_page_70_Picture_9.jpeg)