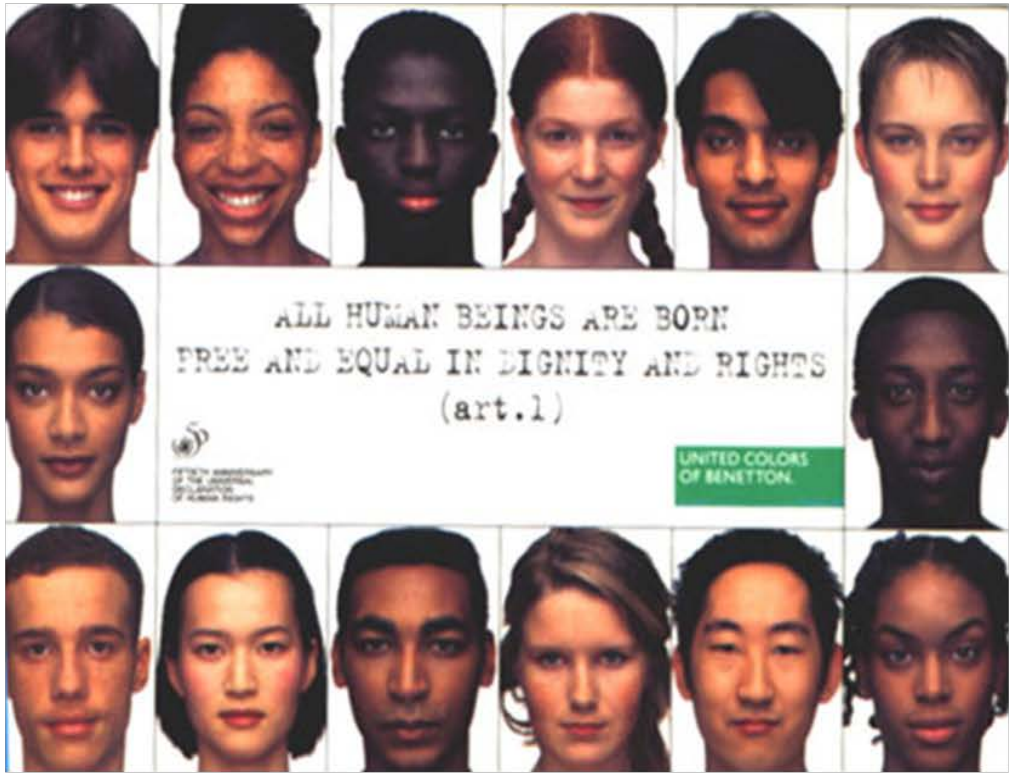
KU LEUVEN

Tools for human molecular diagnostics/Human Genetic diversity

Joris Vermeesch

BeSHG 2021

Why are we different?

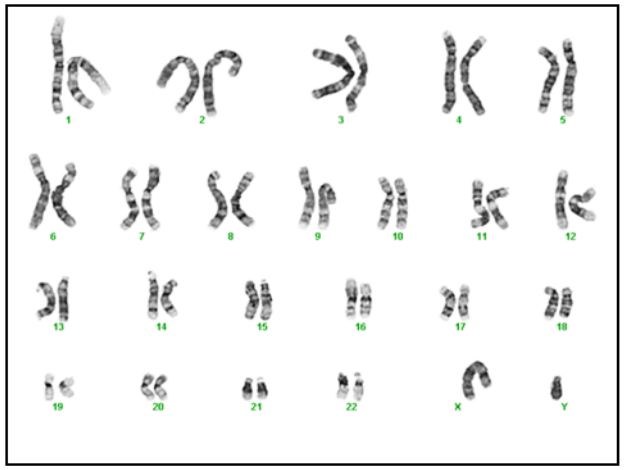


Human Genetic Variation

1. Nature of variation
2. Types of mutations and their consequences
3. Variation in individual genomes
4. Origin and frequency of different types of mutation
5. Consequences for molecular diagnostics of WES/WGS

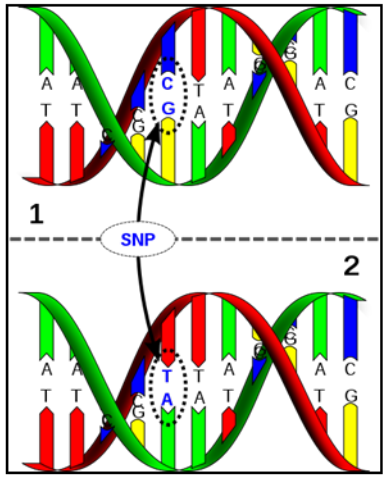
Causes of genetic variation

Chrosomes (1960)



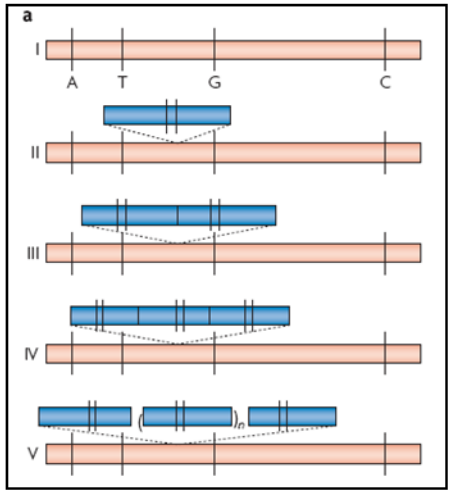
Varianst are rare

Single nucleotide polymorphisms (SNPs) (1980)



Frequent:
- 1 SNP every 1000 bp
- 0.1% difference between 2 human genomes
- 3 Mb difference

Copy number variations (CNVs) (2004)



Very frequent:
- 1000 CNVs/2 individuals
- 0.7% of genome is copy variable between 2 individuals
- 21 Mb difference!

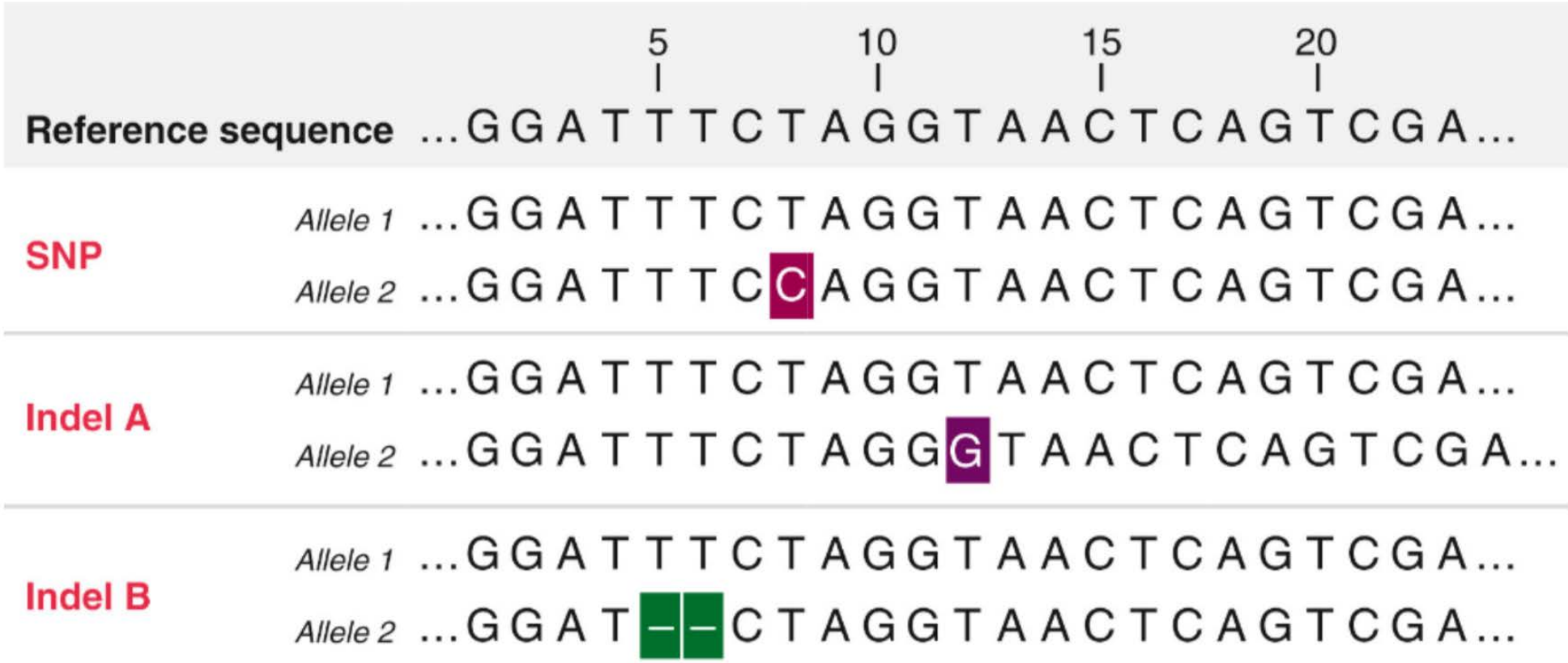
Types of variation and their consequences

- Chromosomes & Copy number variation => see lesson on chromosomes

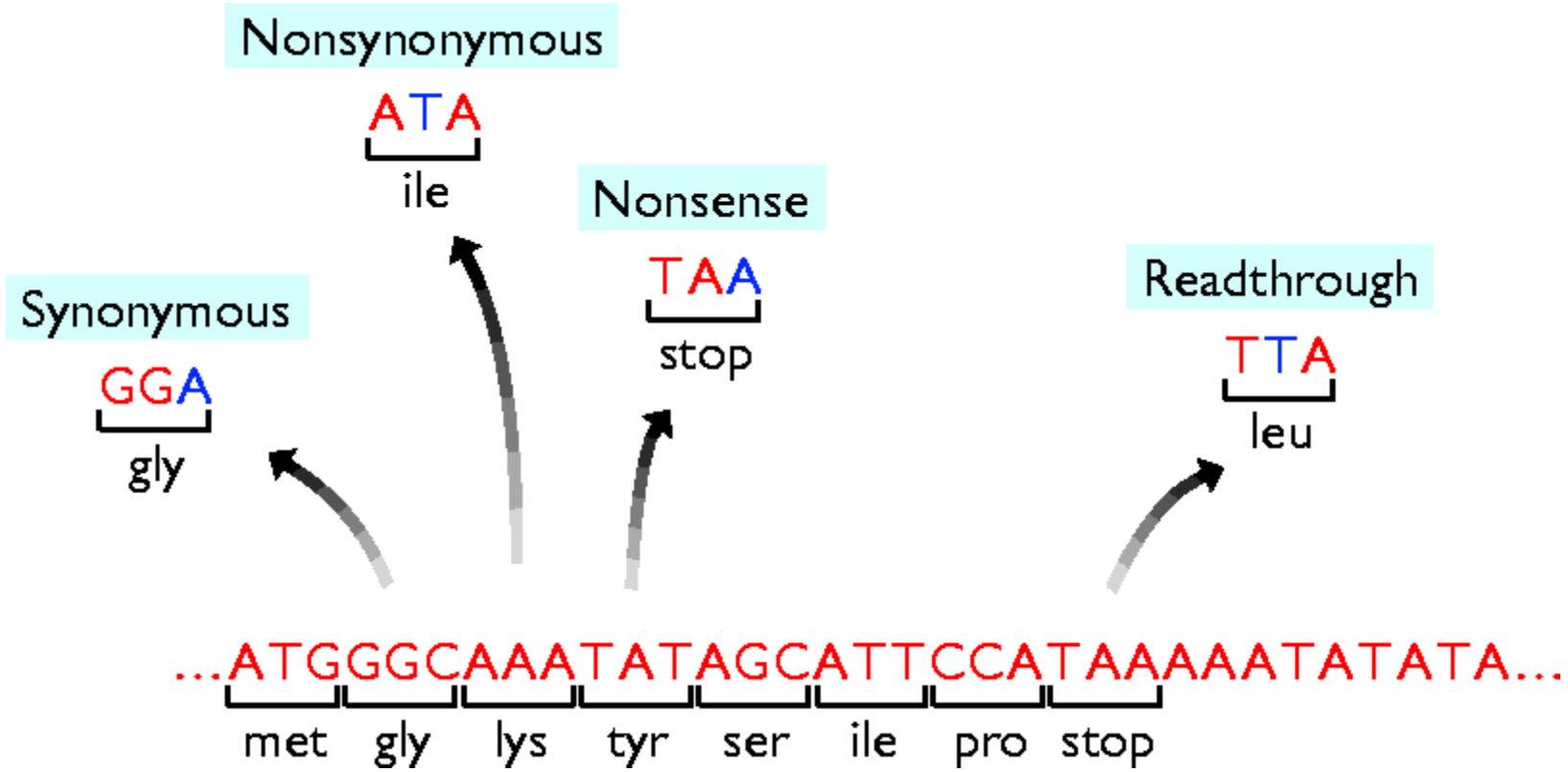
SNPs : Common variation in the genome

Type of Variation	Size Range (approx.)	Basis for the Polymorphism	Number of Alleles
Single nucleotide polymorphisms	1 bp	Substitution of one or another base pair at a particular location in the genome	Usually 2
Insertion/deletions (indels)	1 bp to > 100 bp	*Simple*: Presence or absence of a short segment of DNA 100-1000 bp in length *Microsatellites*: Generally, a 2-, 3-, or 4-nucleotide unit repeated in tandem 5-25 times	*Simple*: 2 *Microsatellites*: typically 5 or more
Copy number variants	10 kb to > 1 Mb	Typically the presence or absence of 200-bp to 1.5-Mb segments of DNA, although tandem duplication of 2, 3, 4, or more copies can also occur	2 or more
Inversions	Few bp to > 1 Mb	A DNA segment present in either of two orientations with respect to the surrounding DNA	2

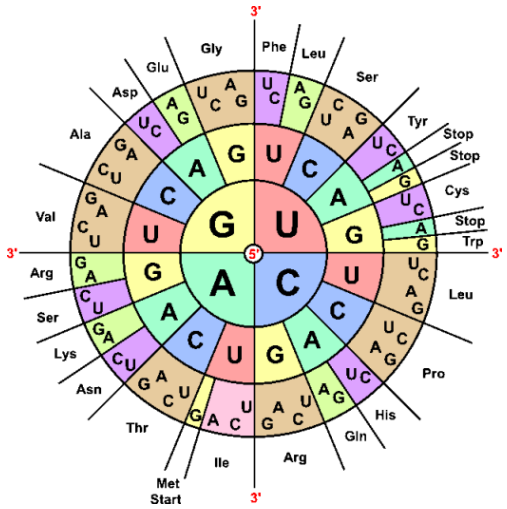
SNPs : Common variation in the genome



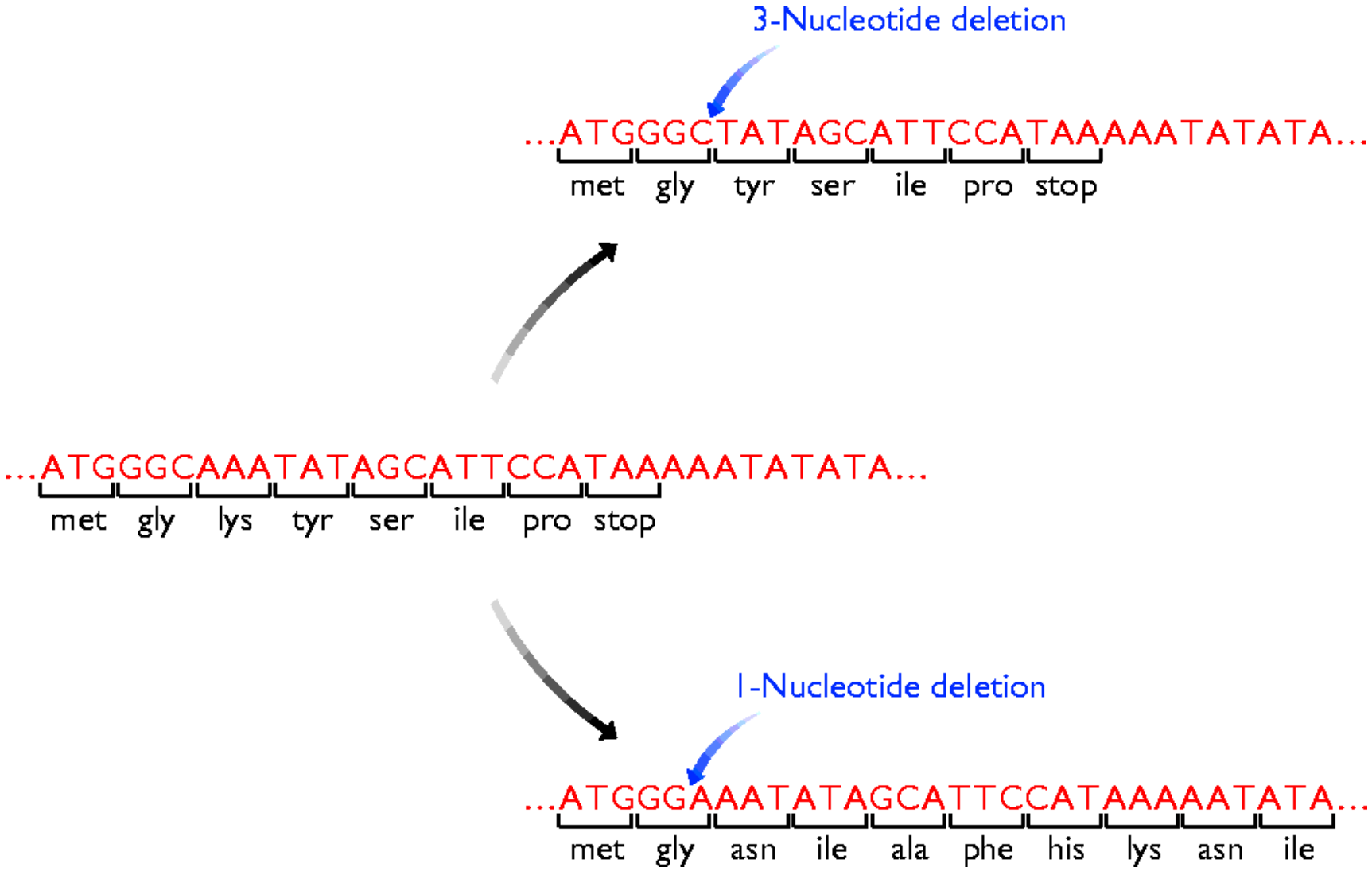
Point mutations (SNV)



First letter	Second letter: U	Second letter: C	Second letter: A	Second letter: G	Third letter
U	UUU Phe	UCU Ser	UAU Tyr	UGU Cys	U
	UUC Phe	UCC Ser	UAC Tyr	UGC Cys	C
	UUA Leu	UCA Ser	UAA Stop	UGA Stop	A
	UUG Leu	UCG Ser	UAG Stop	UGG Trp	G
C	CUU Leu	CCU Pro	CAU His	CGU Arg	U
	CUC Leu	CCC Pro	CAC His	CGC Arg	C
	CUA Leu	CCA Pro	CAA Gln	CGA Arg	A
	CUG Leu	CCG Pro	CAG Gln	CGG Arg	G
A	AUU Ile	ACU Thr	AAU Asn	AGU Ser	U
	AUC Ile	ACC Thr	AAC Asn	AGC Ser	C
	AUA Ile	ACA Thr	AAA Lys	AGA Arg	A
	AUG Met	ACG Thr	AAG Lys	AGG Arg	G
G	GUU Val	GCU Ala	GAU Asp	GGU Gly	U
	GUC Val	GCC Ala	GAC Asp	GGC Gly	C
	GUA Val	GCA Ala	GAA Glu	GGA Gly	A
	GUG Val	GCG Ala	GAG Glu	GGG Gly	G



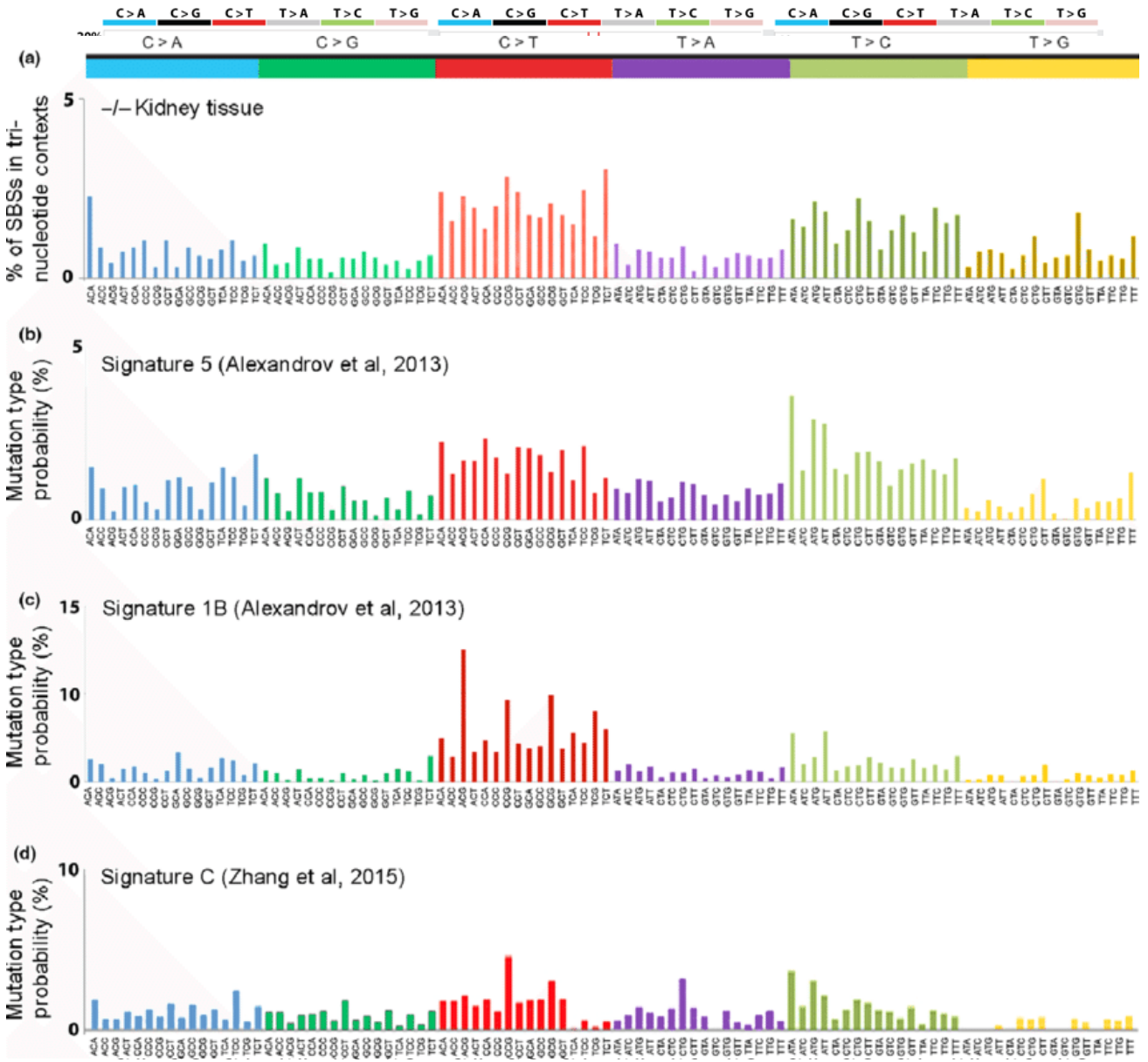
Deletions and insertions (indels)



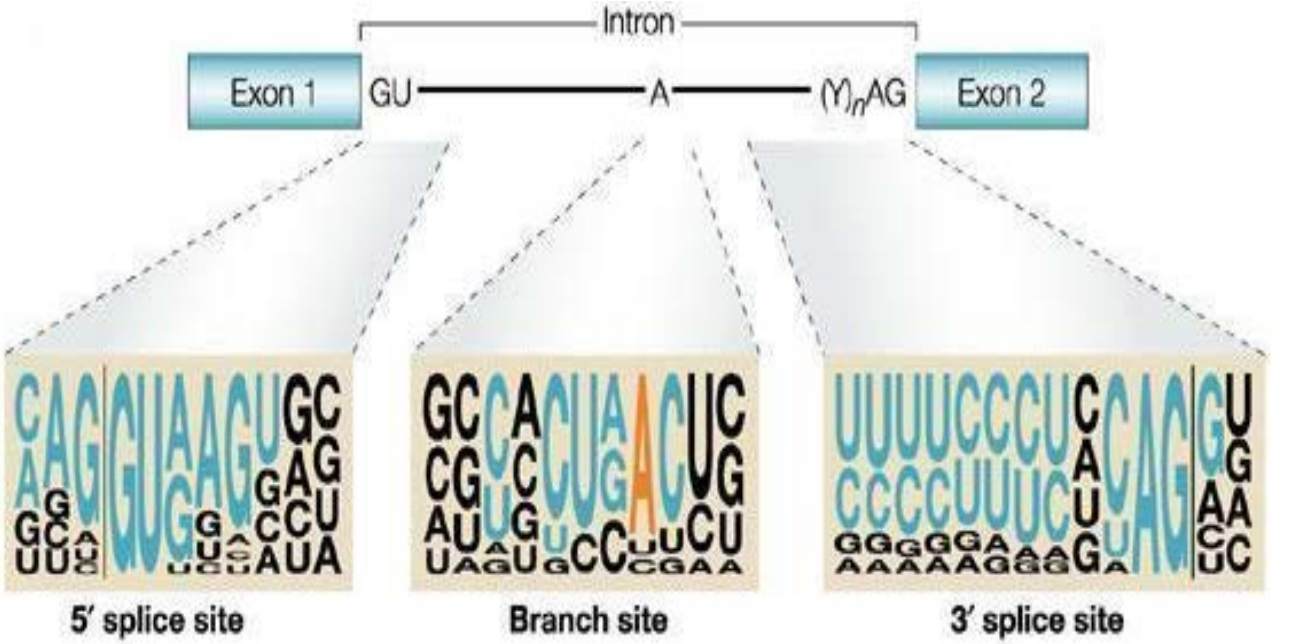
Point mutations origin

- During replication (1 mutation/cell division)
- DNA damage
  - Estimated to be 10000- 1M nucleotides are damaged/human/day
  - Spontaneous chemical processes: e.g. Depurination, Demethylation, Deamination
  - Chemical mutagens (natural or otherwise)
  - Ionizing and UV radiation
- DNA damage is repaired, but some remain.

Mutational signatures

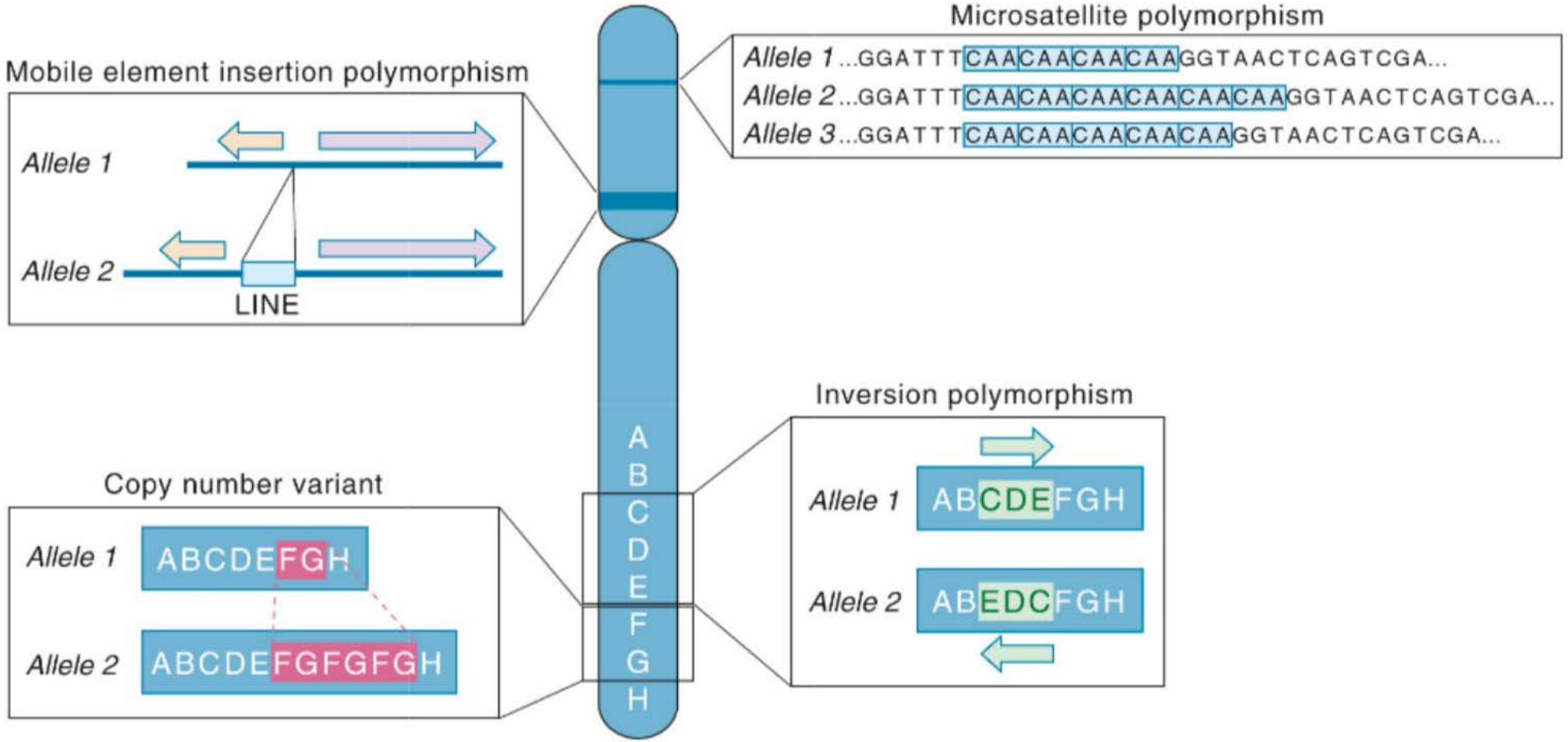


Splice consensus signals

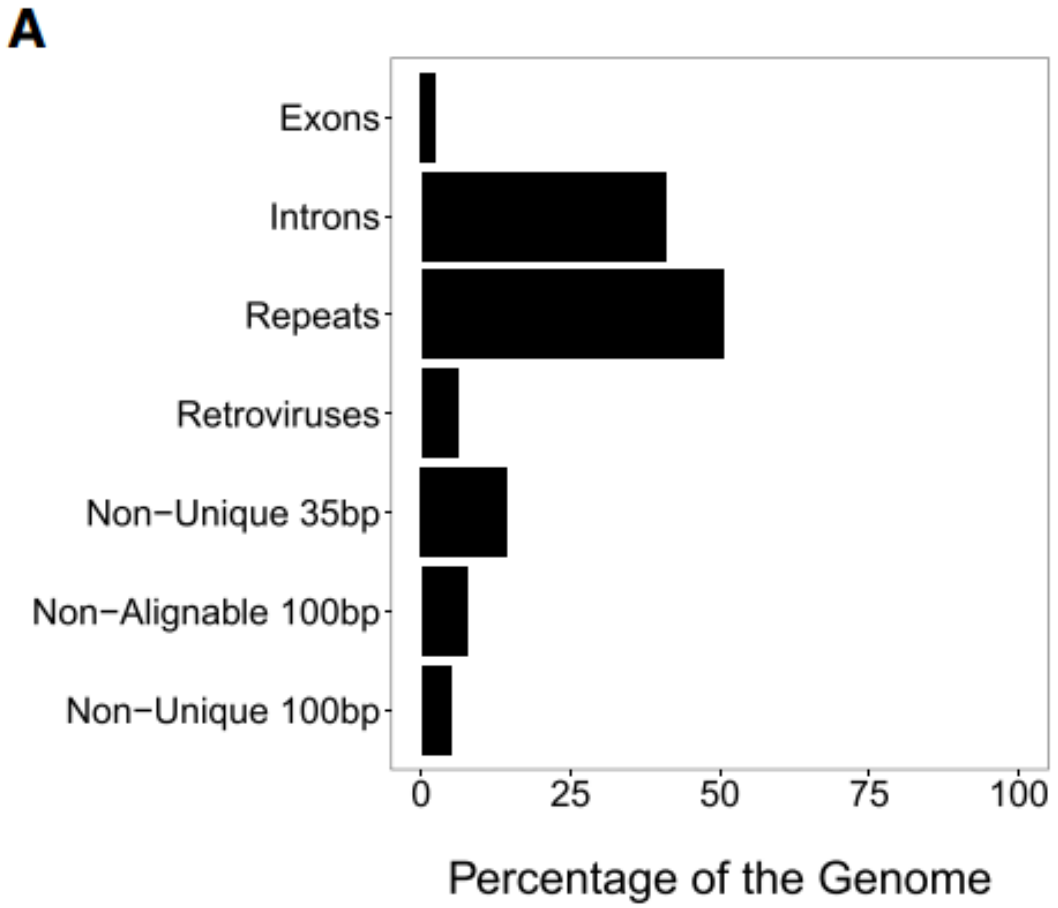


Figuur 1

Nature Reviews | Genetics

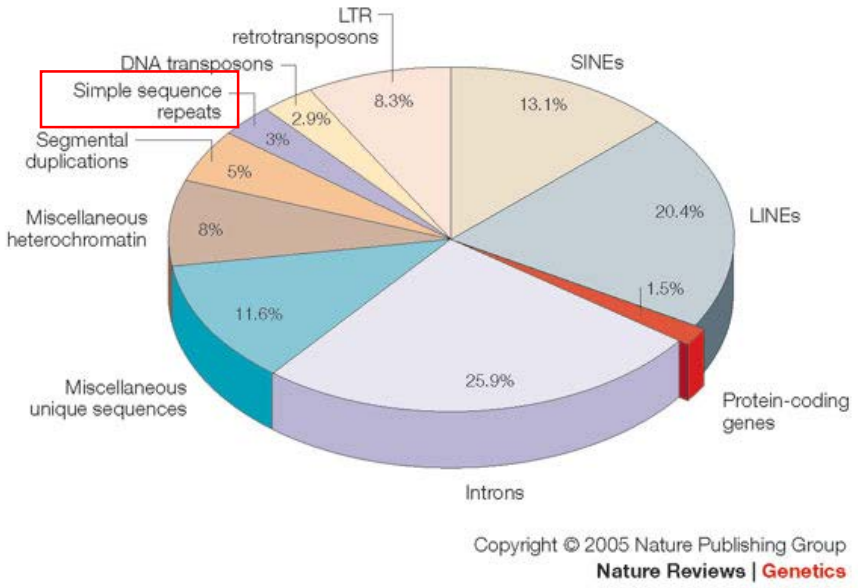

Microsatellite polymorphism
Allele 1 ...GGATTTCAACAACAACAACAAGGTAACTCAGTCGA...
Allele 2 ...GGATTTCAACAACAACAACAACAACAAGGTAACTCAGTCGA...
Allele 3 ...GGATTTCAACAACAACAACAACAAGGTAACTCAGTCGA...
Mobile element insertion polymorphism
Allele 1
Allele 2
LINE
A
B
C
D
E
F
G
H
Inversion polymorphism
Allele 1
ABCDEFGH
Allele 2
ABEDCFGH
Copy number variant
Allele 1
ABCDEFGH
Allele 2
ABCDEFGFGFGH

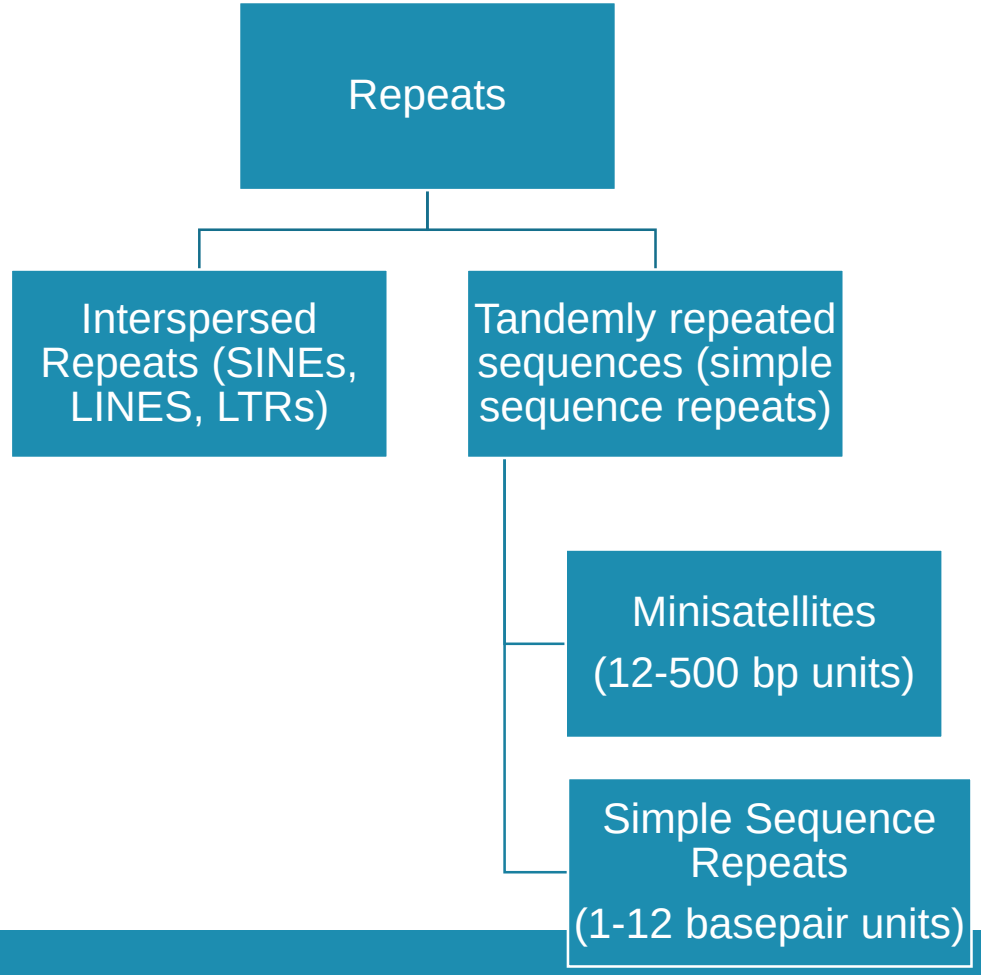
Het humane genoom bevat 50% repeats



Goldfeder et al. (2016). Medical implications of technical accuracy in genome sequencing. Genome Medicine (2016) 8:24

Het humane genoom bevat 50% repeats





Short tandem repeat (STR)

- Short tandem repeat = Microsattelites = Variable number of tandem repeats = simple sequence repeats

- They have specific unit: e.g. CGG

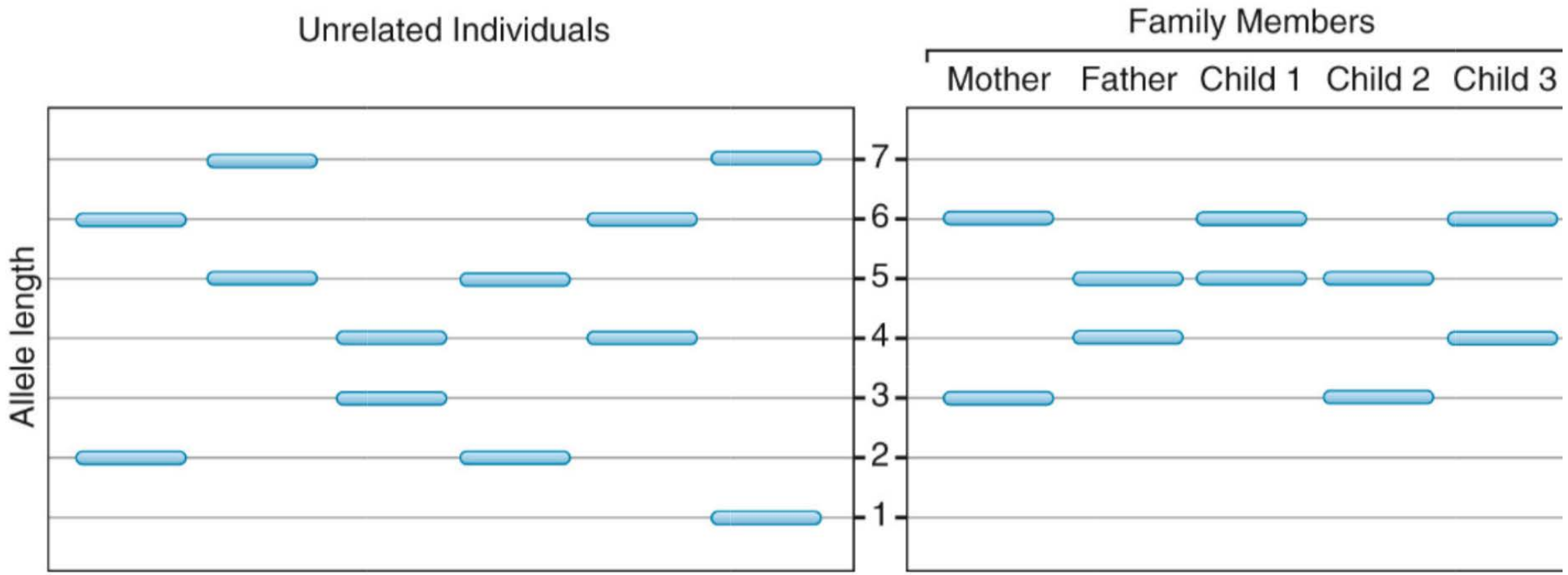
- That is repeated:

  CGGCGGCGGCGGCGGCGG

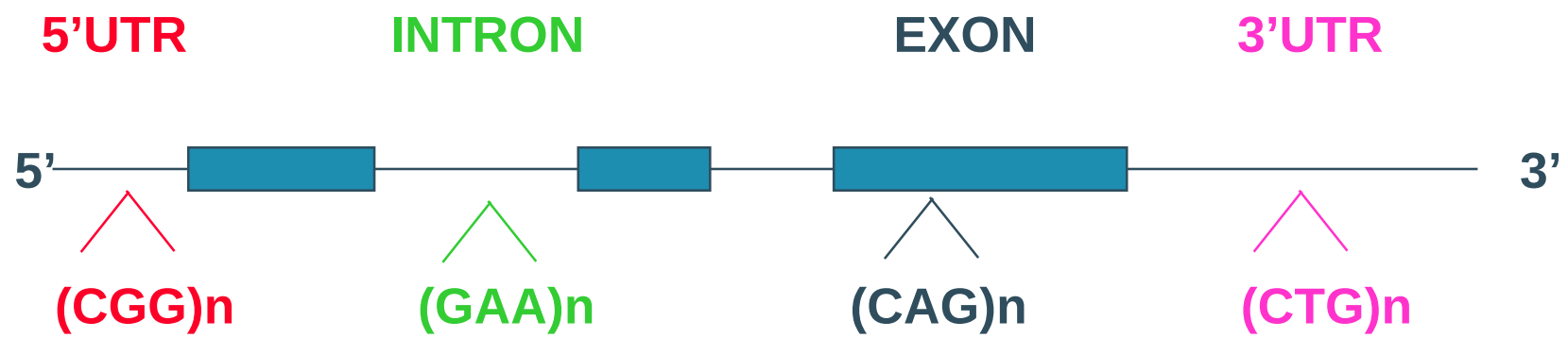
- **± 1 miljoen STRs** in the human genome

- Tandem repeats can have a big impact on phenotype

A schematic of a hypothetical microsatellite marker in human DNA.



Tandem repeats in genes are associated with disorders characterized by anticipation



Fragiele X

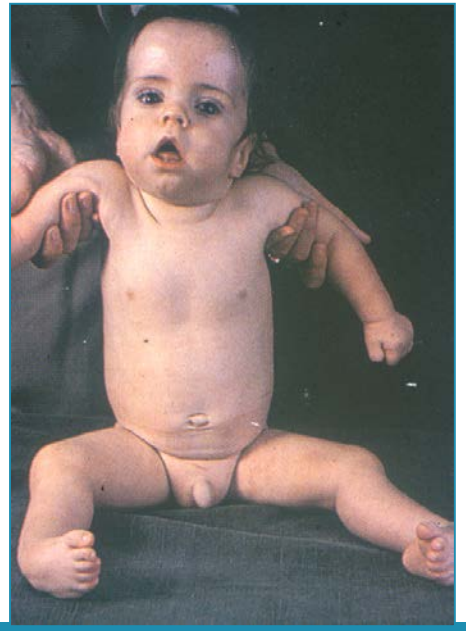
Friedreich ataxie

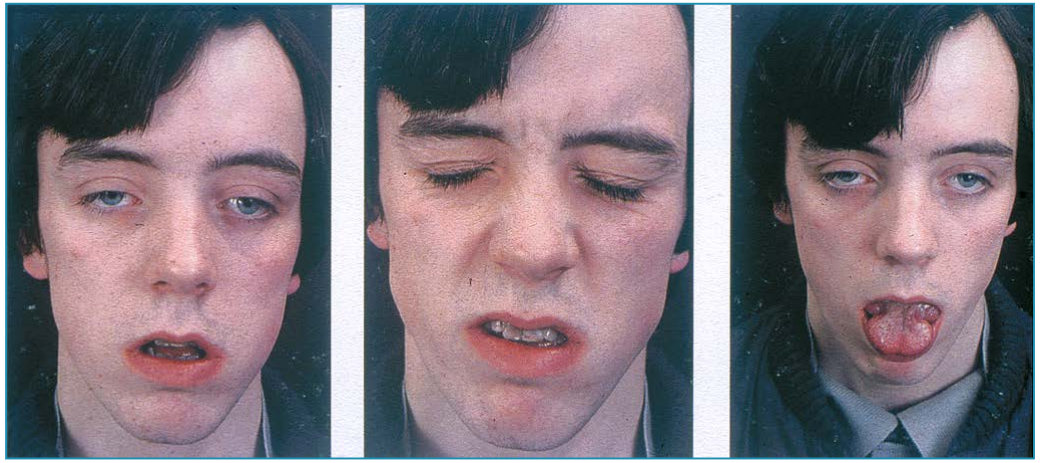
Huntington
Kennedy
SCA1, 2, 3, 6, 7
DRPLA

Steinert
SCA8

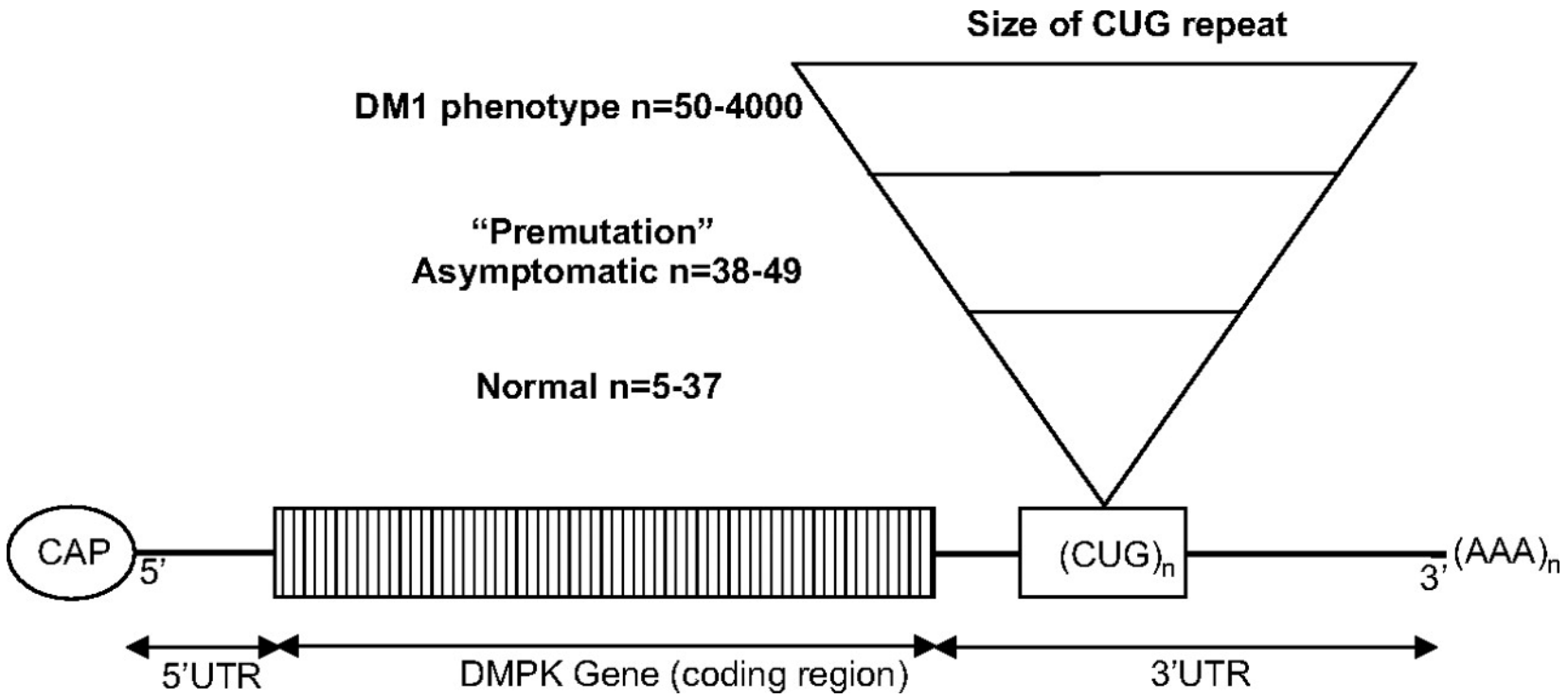
Myotonic dystrophy (Steinert disease) (as an example)

- *Autosomal dominant*
- *Trinucleotide repeat expansion*





DMPK pre-mRNA with relationship between CUG repeat size and phenotype.

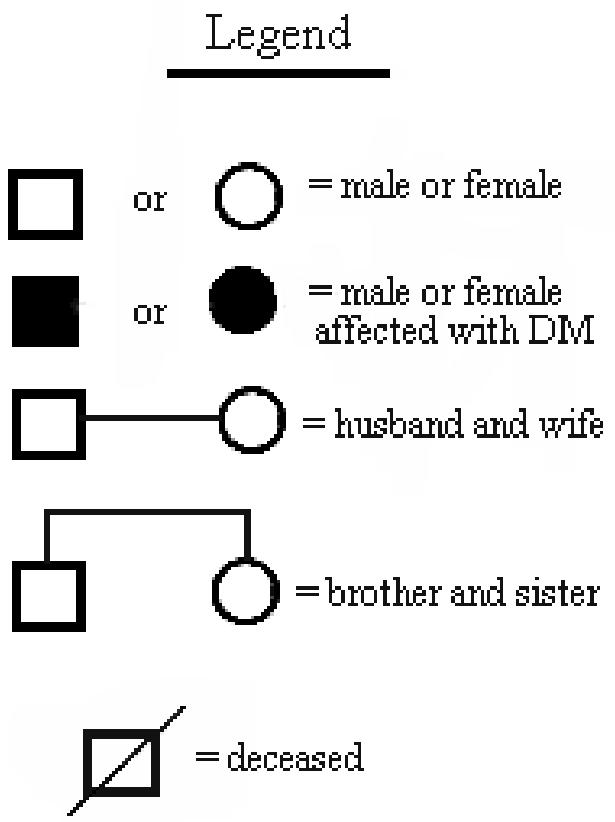


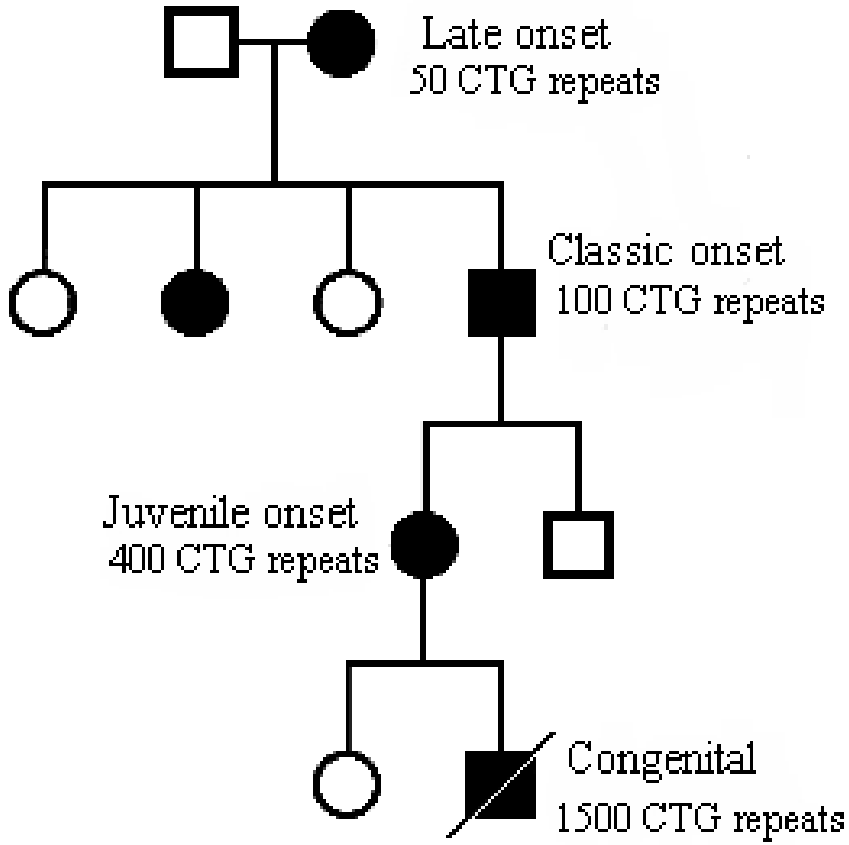
Chris Turner, and David Hilton-Jones J Neurol Neurosurg Psychiatry 2010;81:358-367

Myotonic dystrophy (Steinert disease)
Hereditary aspects

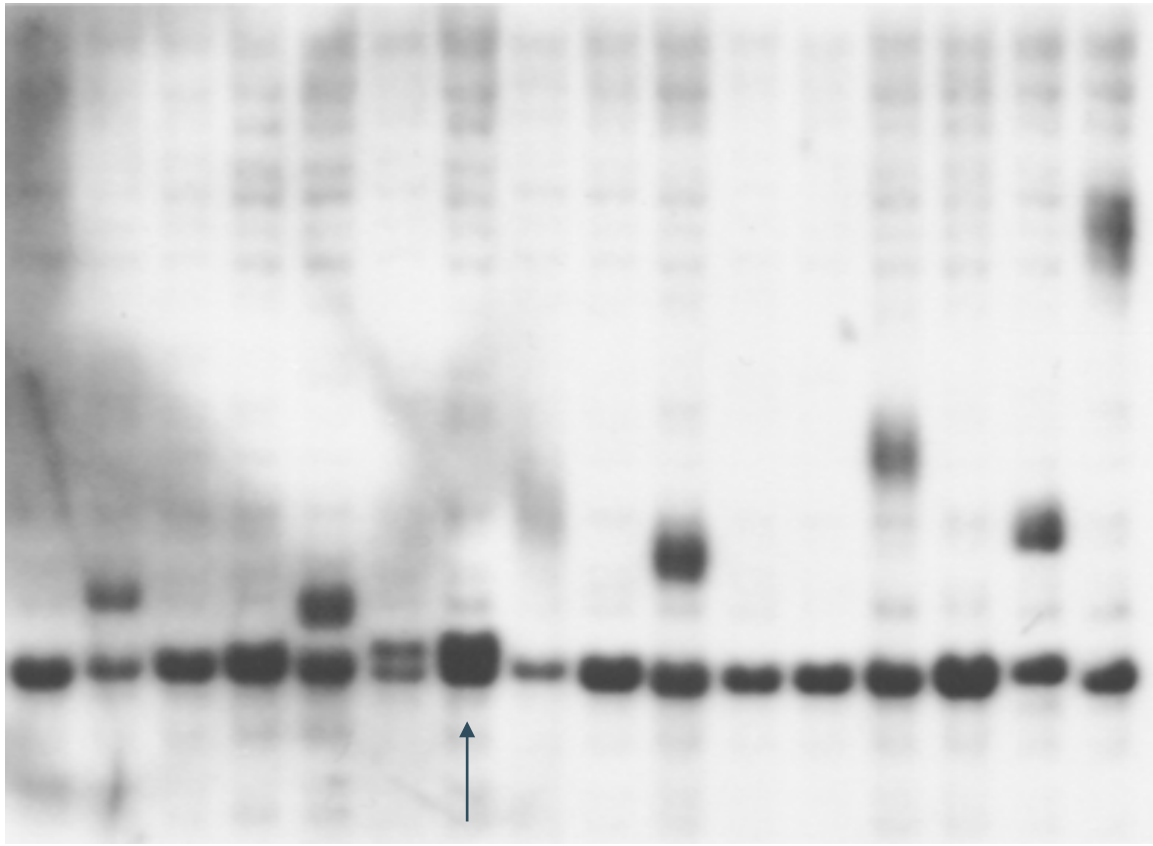
- Anticipation
  - Increasing severity and successive generations
- Maternal transmission for large expansions
- Often paternal transmission in case of smaller expansions.

DM: anticipation





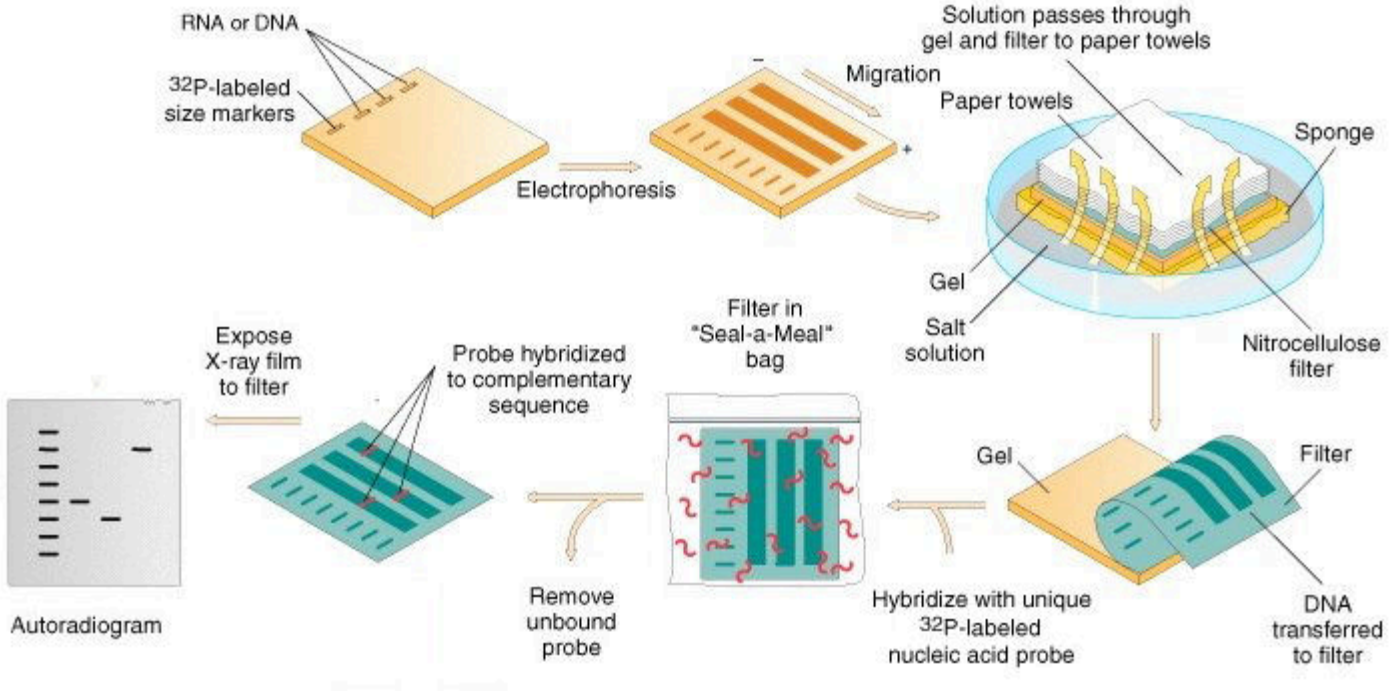
Myotonic dystrophy (Steinert disease): detection



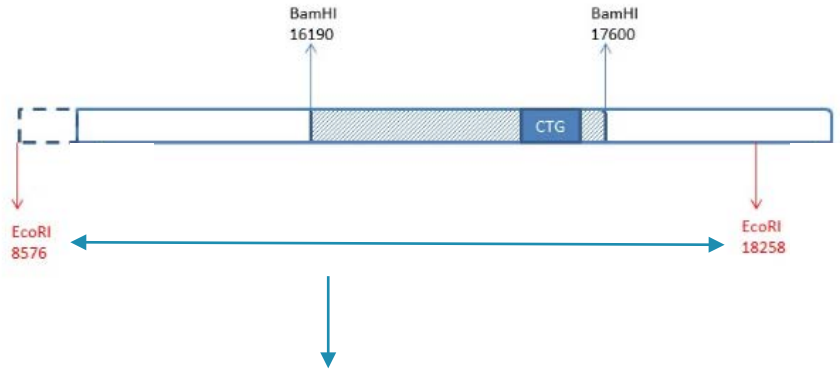
n = 35

Southern blot om lange expansies te bepalen:

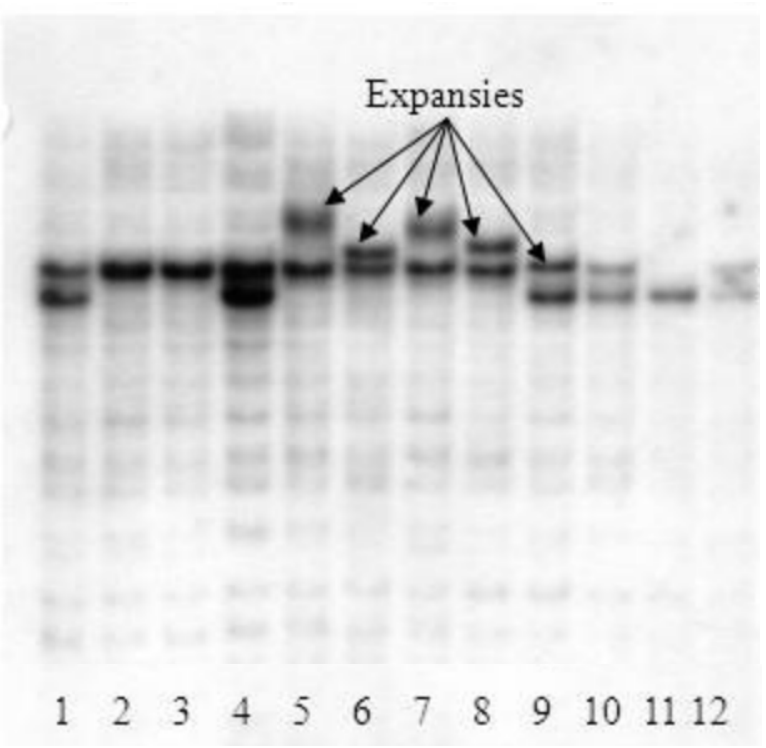
1. Knippen van humaan genoom met restrictie enzymes



Southern blot to detect large expansions

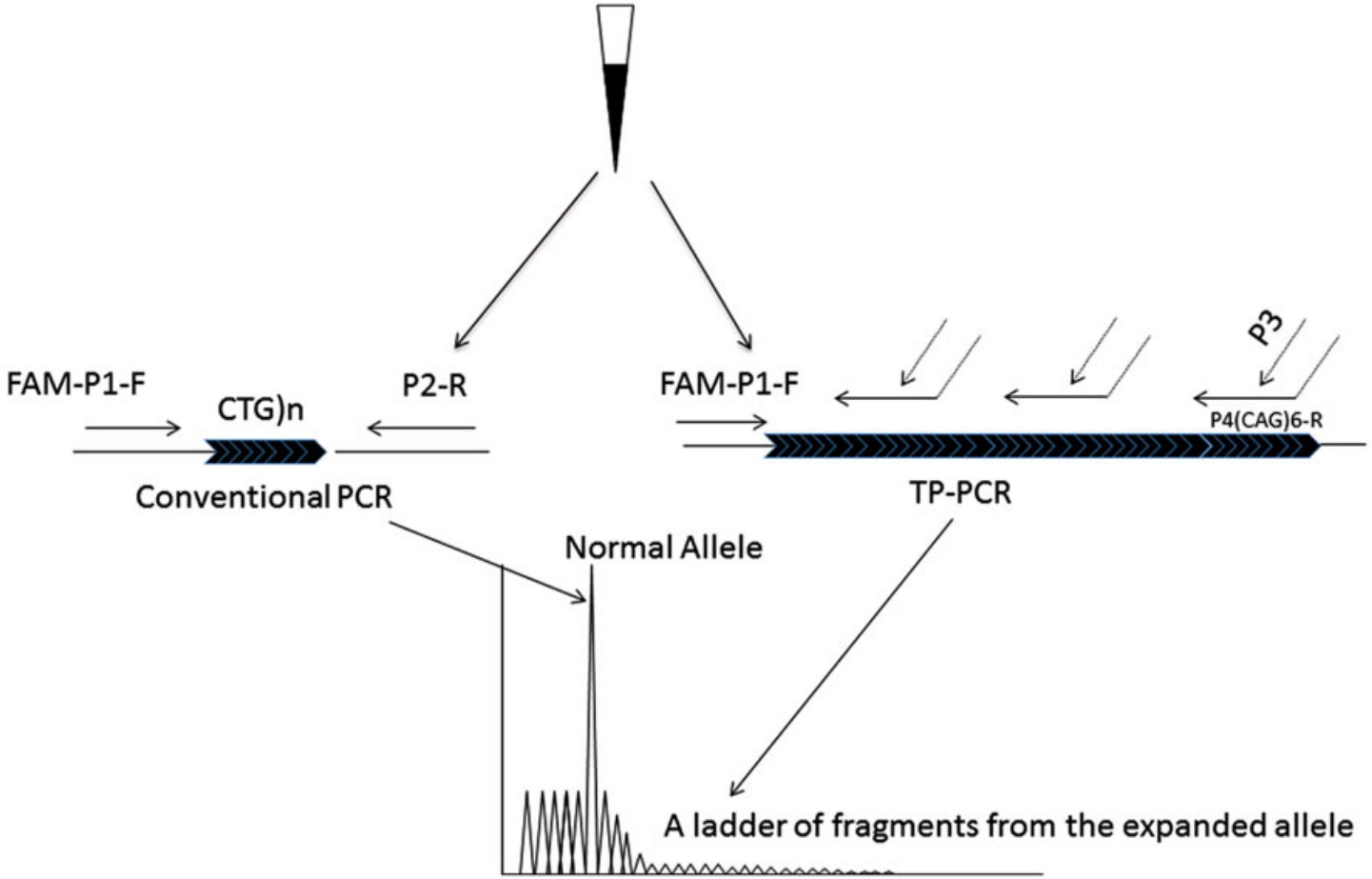


9682 bp



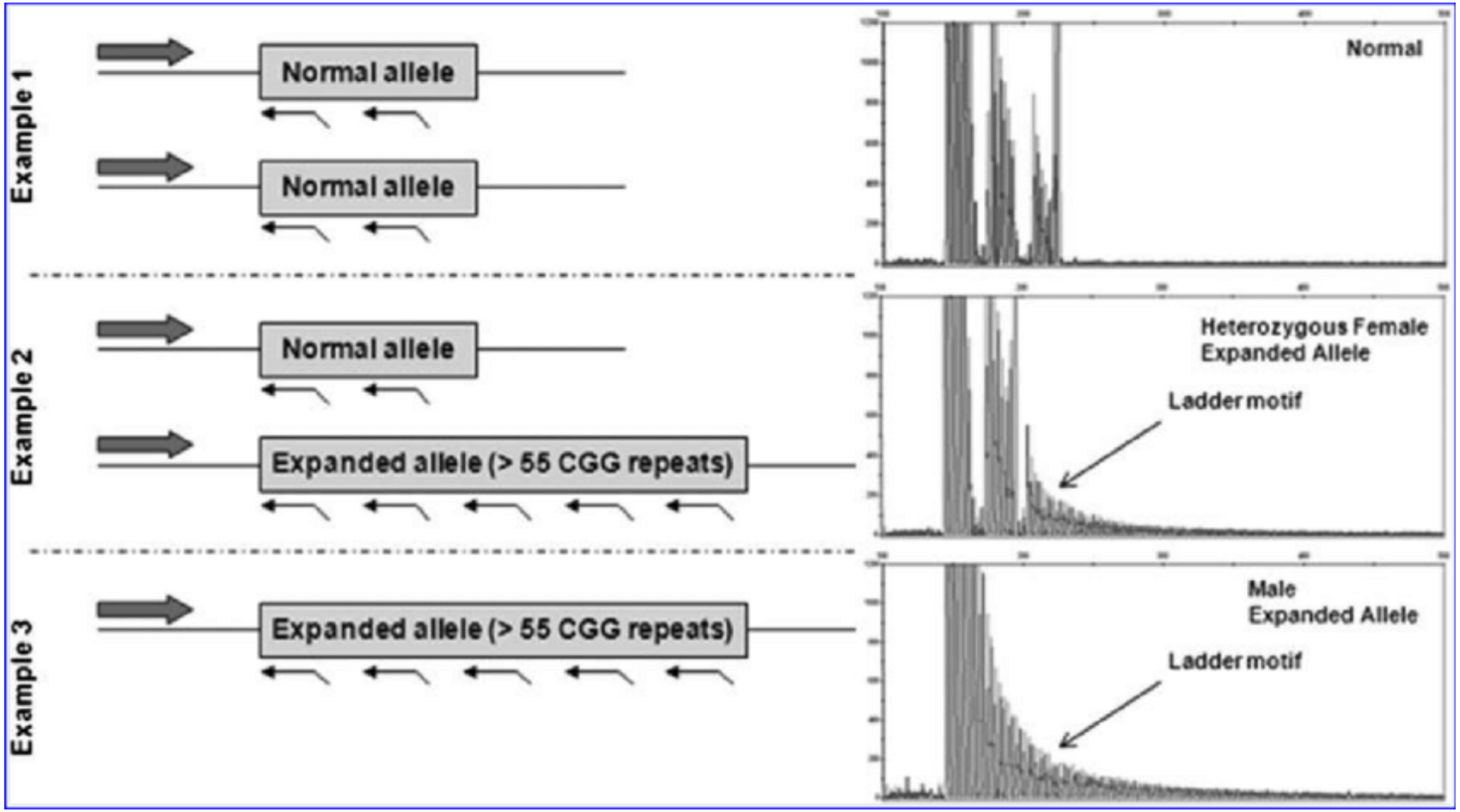
-> Fragment can be recognized via probe
-> Fragment will be larger with larger expansions

Long expansions can be detected by Triplet primed-PCR (TP-PCR)



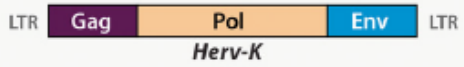
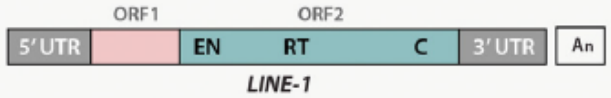
Kan de lengte van lange expansies niet bepalen

Triplet primed PCR

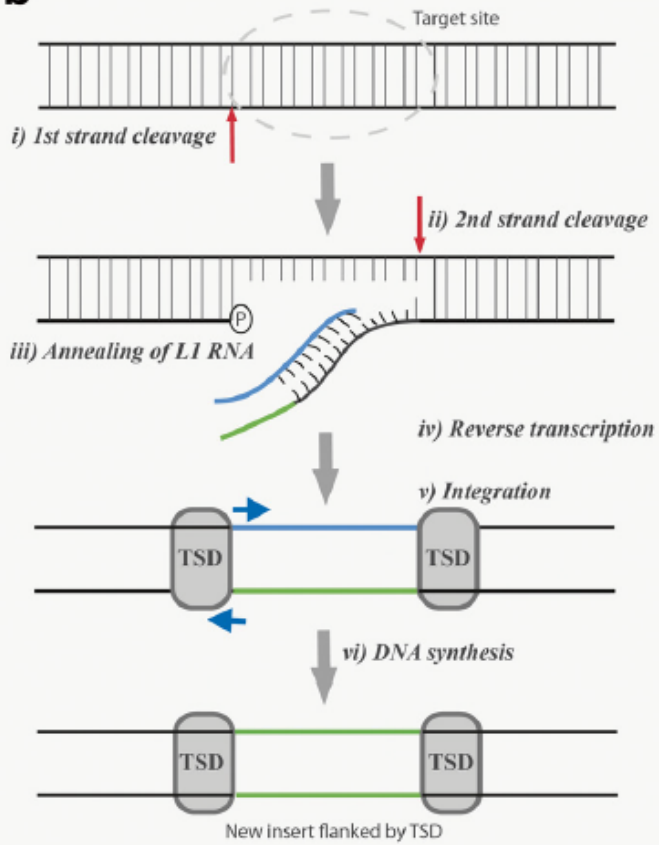


Active mobile elements in the human genome

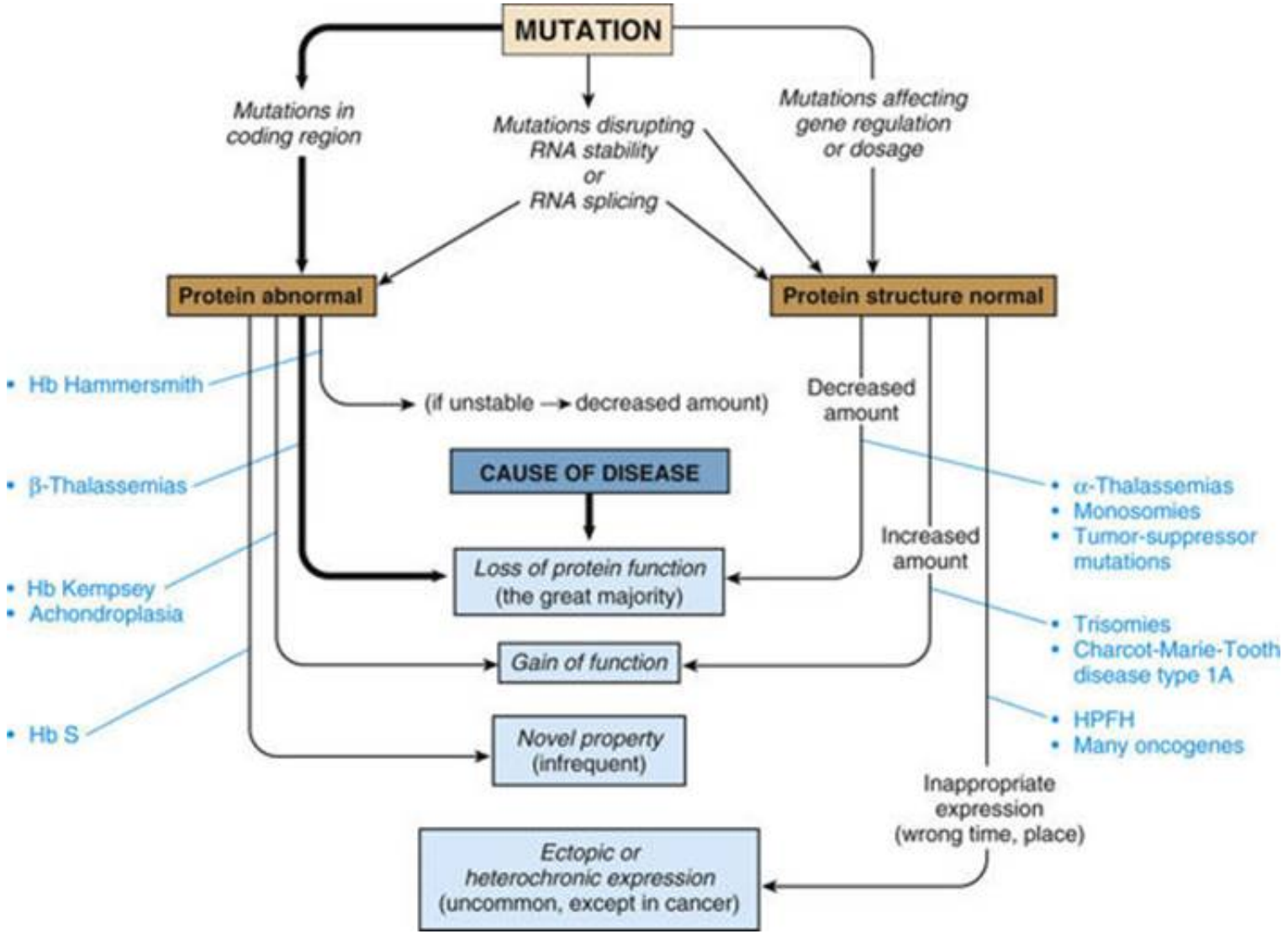
a

Mobile element structure	HGR	Length	Remarks
DNA transposons: 	2-3%	1.4 kb	Transposes from less than 100 kb to distant sites from original site
Retrotransposons (autonomous): 	7-9%	±1.4 kb	Retrovirus like structure with defective envelope gene, reinserts in the same genome from which they come
	17-19%	6 kb	Only autonomously active mobile elements in primates and humans
Non-autonomous: 		0.3 kb	Alu insertions accounts for over 20 cases of human genetic diseases
	11-13%	±1.5 kb	SVA insertions occurs at high frequency, so far 3 cases of human disease reported
		Variable	Arises by revere transcription of cellular mRNA & integration of cDNA in the genome

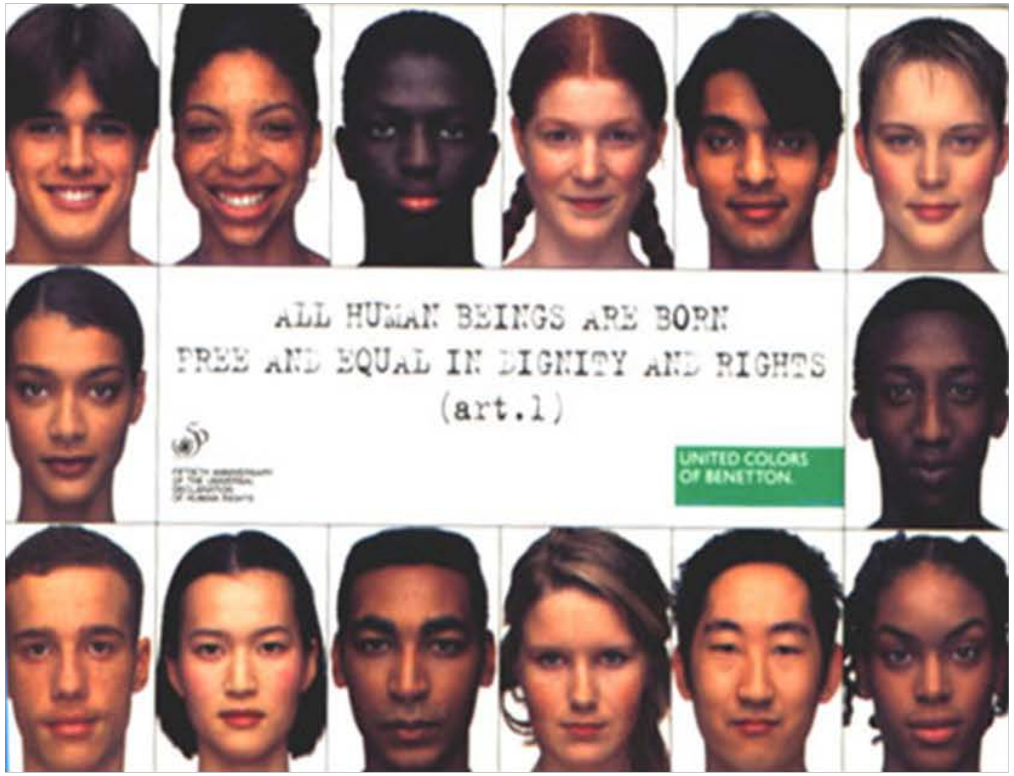
b



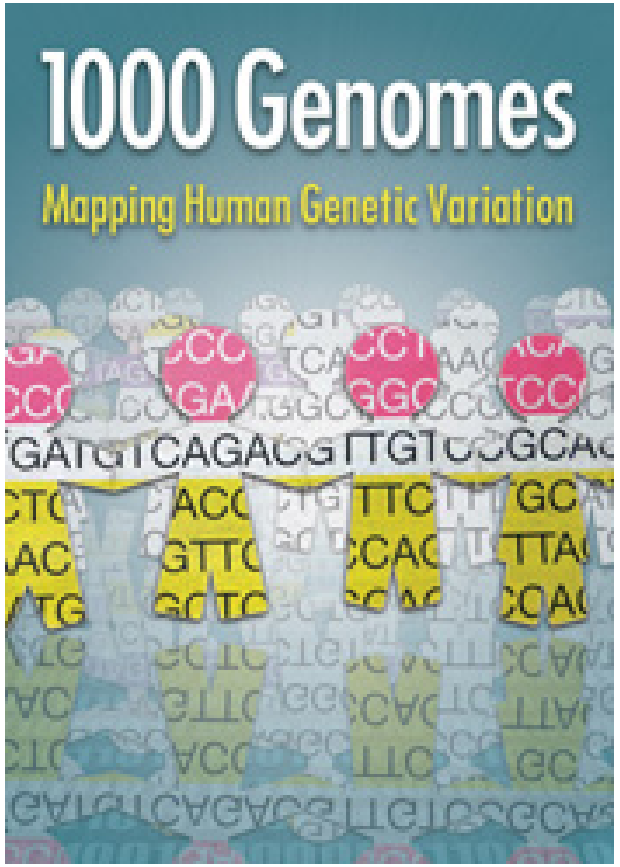
Sargurupremraj and Wjst Respiratory Research 2013, 14:99


MUTATION
Mutations in coding region
Mutations disrupting RNA stability or RNA splicing
Mutations affecting gene regulation or dosage
Protein abnormal
Protein structure normal
• Hb Hammersmith
(if unstable → decreased amount)
Decreased amount
CAUSE OF DISEASE
• β-Thalassemias
• α-Thalassemias
• Monosomies
• Tumor-suppressor mutations
Increased amount
Loss of protein function (the great majority)
• Hb Kempsey
• Achondroplasia
• Trisomies
• Charcot-Marie-Tooth disease type 1A
Gain of function
• Hb S
Novel property (infrequent)
• HPFH
• Many oncogenes
Inappropriate expression (wrong time, place)
Ectopic or heterochronic expression (uncommon, except in cancer)

Variation in individual genomes



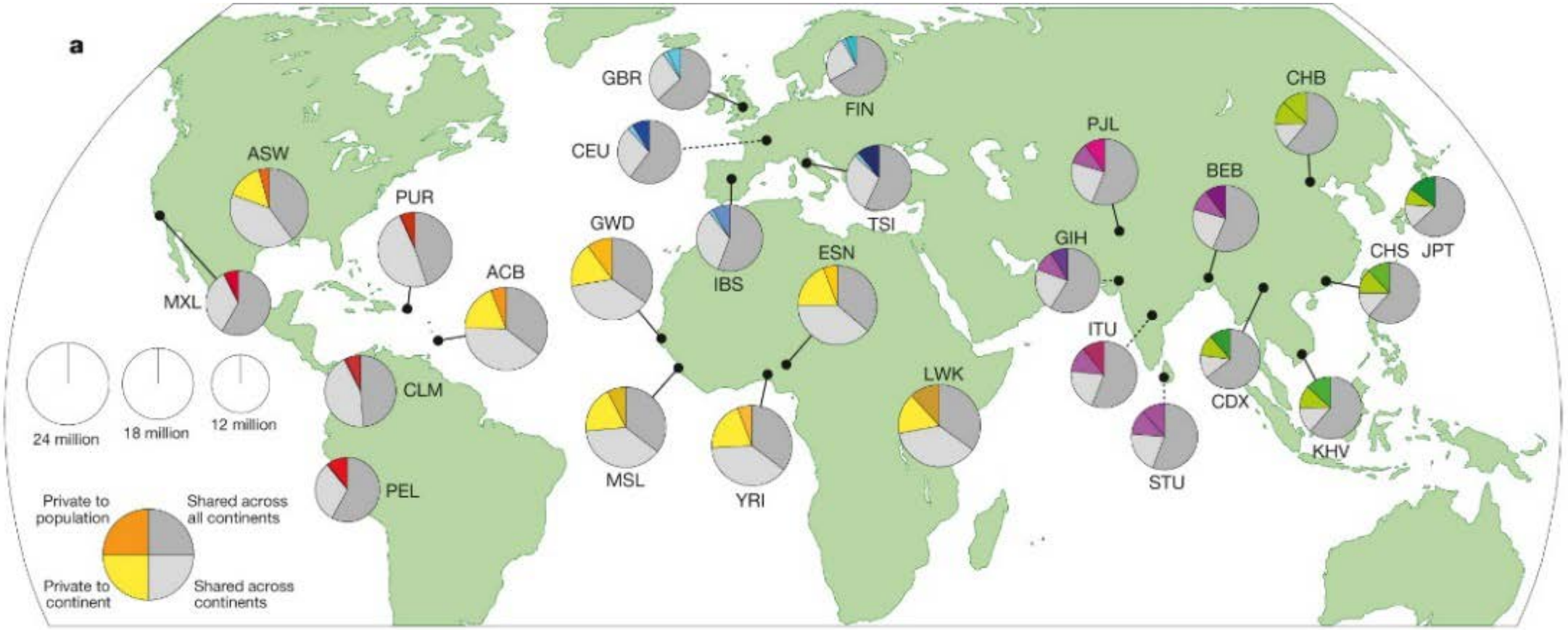
1000 genome project/resource



- Comprehensive description of common human genetic variation
- Latest report: genomes of 2,504 individuals from 26 populations using a combination of low-coverage whole-genome sequencing, deep exome sequencing, and dense microarray genotyping.
- Results:
  - over **88 million variants** (84.7 million single nucleotide polymorphisms (SNPs)
  - **3.6 million short insertions/deletions** (indels), and 60,000 structural variants), all phased onto high-quality haplotypes.
  - This resource includes >99% of SNP variants with a frequency of >1% for a variety of ancestries.

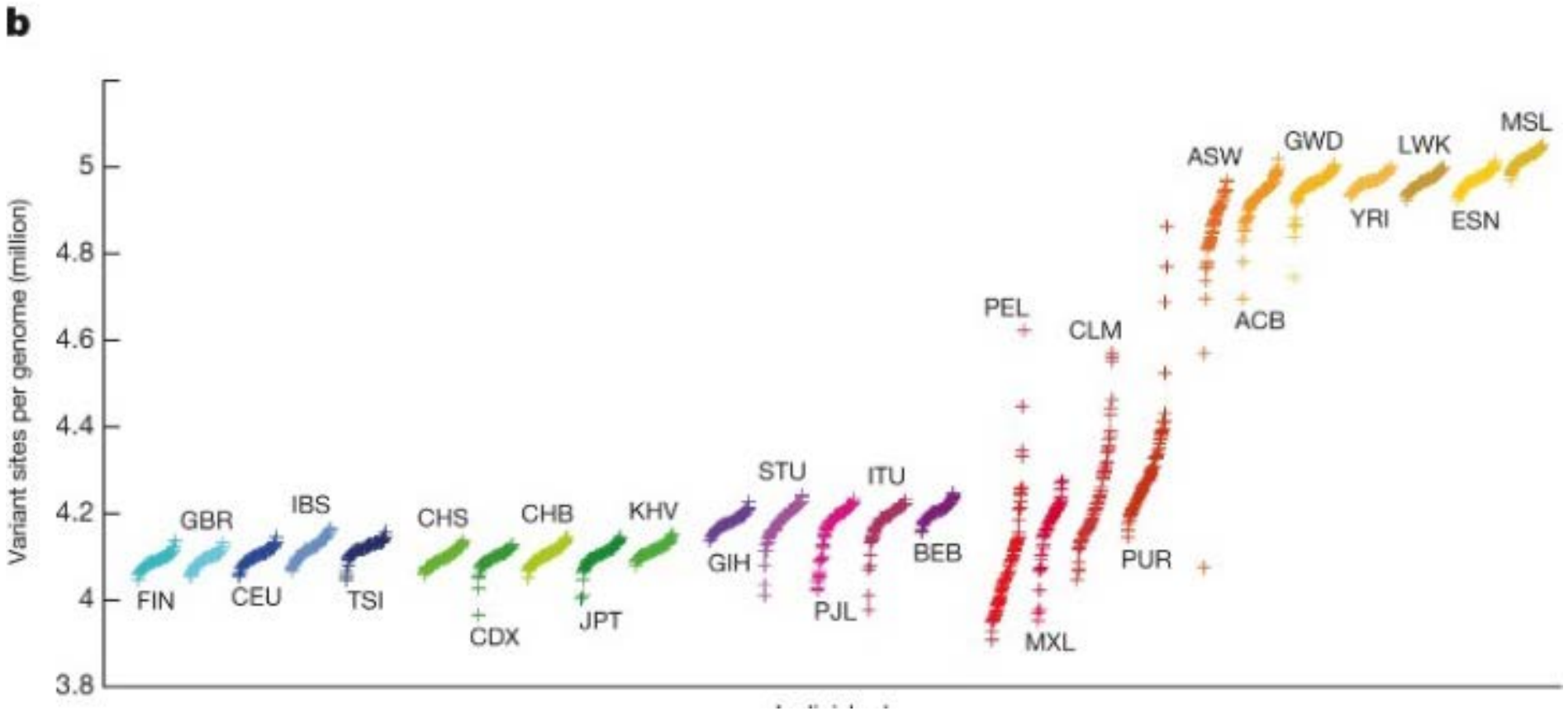
A Auton *et al*. *Nature* **526**, 68-74 (2015) doi:10.1038/nature15393

SNP variation/population



A Auton *et al*. *Nature* **526**, 68-74 (2015)
doi:10.1038/nature15393

SNP variation/population

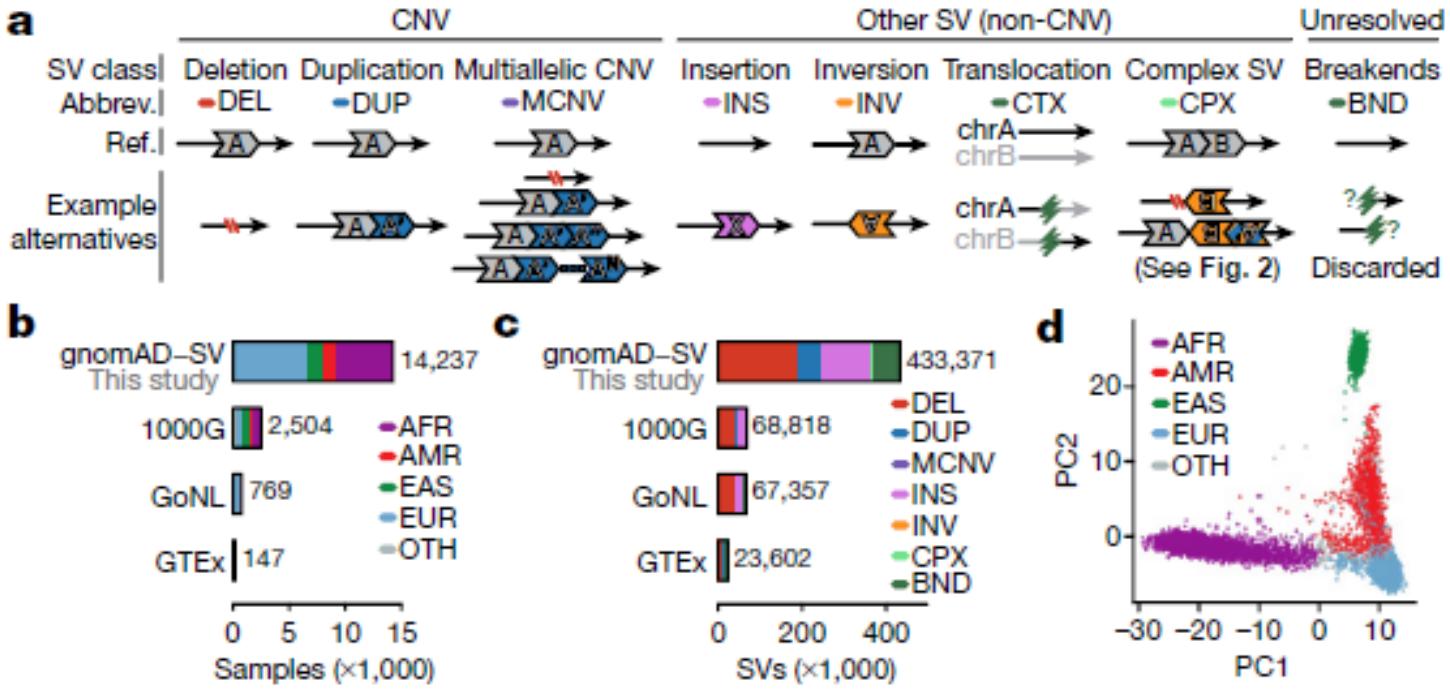


A Auton *et al*. *Nature* **526**, 68-74 (2015) doi:10.1038/nature15393

Structural variation

genome aggregation database or GnomAD

Based on short read sequencing in 14290 genomes



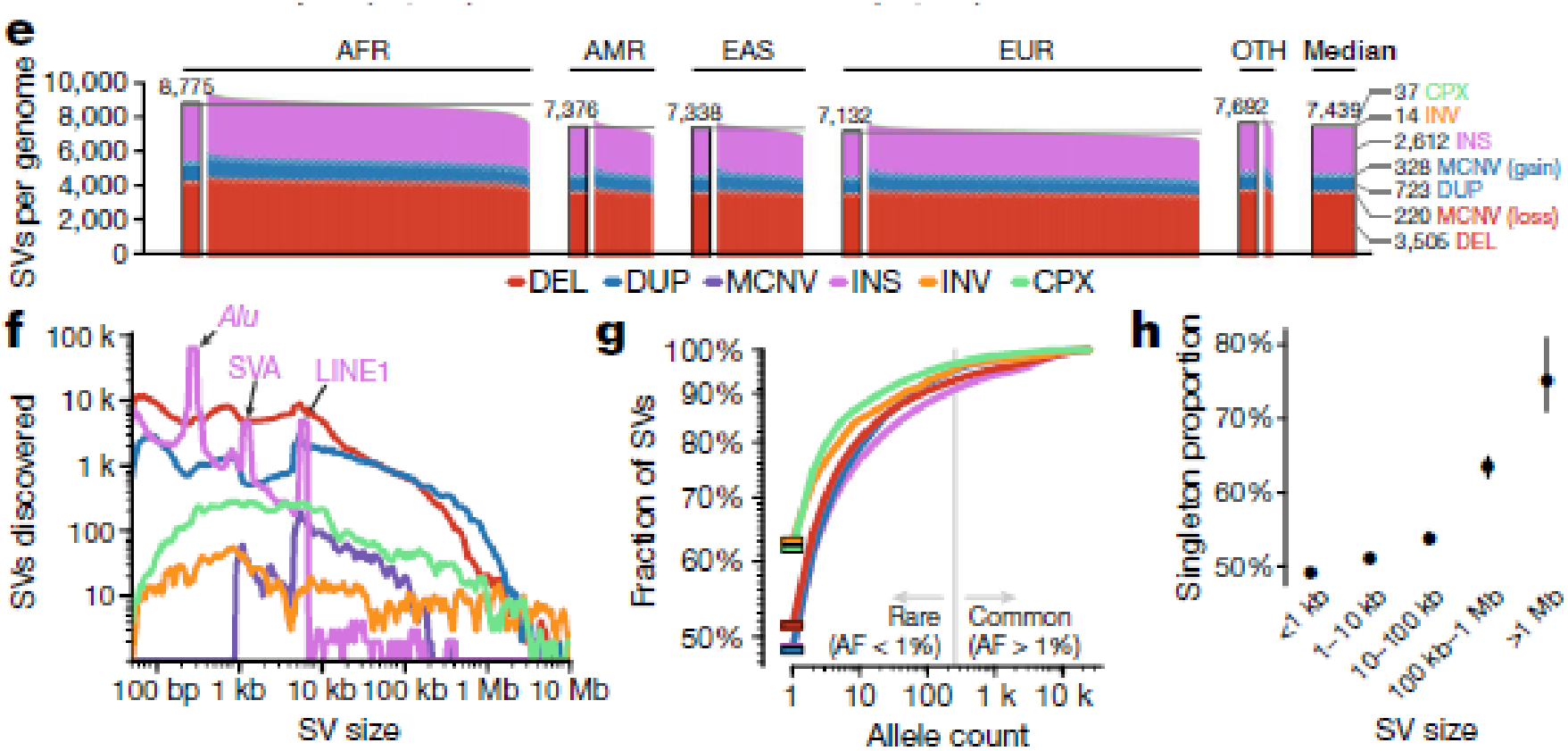
Article

A structural variation reference for medical and population genetics

https://doi.org/10.1038/s41586-020-2287-8
Received: 2 March 2019
Accepted: 31 March 2020
Published online: 27 May 2020
Open access
Check for updates

Ryan L. Collins[1,2,3,157], Harrison Brand[1,2,4,157], Konrad J. Karczewski[1,5], Xuefang Zhao[1,2,4], Jessica Alföldi[1,5], Laurent C. Francioli[1,5,6], Amit V. Khera[1,3], Chelsea Lowther[1,2,4], Laura D. Gauthier[1,7], Harold Wang[1,2], Nicholas A. Watts[1,5], Matthew Solomonson[1,5], Anne O'Donnell-Luria[1,5], Alexander Baumann[7], Ruchi Munshi[7], Mark Walker[1,7], Christopher W. Whelan[7], Yongqing Huang[7], Ted Brookings[7], Ted Sharpe[7], Matthew R. Stone[1,2], Elise Valkanas[1,2,3], Jack Fu[1,2,4], Grace Tiao[1,5], Kristen M. Laricchia[1,5], Valentin Ruano-Rubio[7], Christine Stevens[1], Namrata Gupta[1], Caroline Cusick[1], Lauren Margolin[1], Genome Aggregation Database Production Team*, Genome Aggregation Database Consortium*, Kent D. Taylor[8], Henry J. Lin[8], Stephen S. Rich[9], Wendy S. Post[10], Yii-Der Ida Chen[8], Jerome I. Rotter[8], Chad Nusbaum[1,154], Anthony Philippakis[7], Eric Lander[1,11,12], Stacey Gabriel[1], Benjamin M. Neale[1,2,5,13], Sekar Kathiresan[1,2,6,14], Mark J. Daly[1,2,5,13], Eric Banks[7], Daniel G. MacArthur[1,2,5,6,155,156] & Michael E. Talkowski[1,2,4,13]✉

Structural variation characteristics



Mapping full spectrum of structural variation

- Human Structural variation consortium
- Comprehensive structural variation analysis with a multitude of techniques.
- 3 parent-child trios (Han, Puerto Rican, Yoruban)

Table 1 Summary of sequencing statistics

	Avg. seq. coverage	Avg. frag. length	Physical coverage
Pacific Biosciences	39.6 (child) 20.03 (parent)	8165 (child) 9619 (parent)	39.6
Oxford Nanopore	18.9 (HG00733)	11,993	18.9
Illumina short insert	74.5	694	171
Illumina lWGS	3	3475	159
Illumina 7 kb JMP	1.1	6973.2	39.2
10X Chromium	82.4	90,098	53.9
Bionano Genomics	N/A	2.81E + 05	116.7
Tru-Seq SLR	3.47	4900	3.47
Strand-seq	N/A	N/A	5.87
Hi-C	19.49	1.03E + 07	N/A
Total	223.56		607.08

Physical coverage is given for Illumina short insert, lWGS, 7 kb JMP, 10X Chromium physical coverage is estimated read cloud coverage
For Hi-C, fragment length is the distance between two read ends for intra-chromosome read pairs

ARTICLE

https://doi.org/10.1038/s41467-018-08148-z OPEN

Multi-platform discovery of haplotype-resolved structural variation in human genomes

Mark J.P. Chaisson et al.#

Per genome variation

(3-7x more than known from short read sequencing)

818000 indels (<50bp)

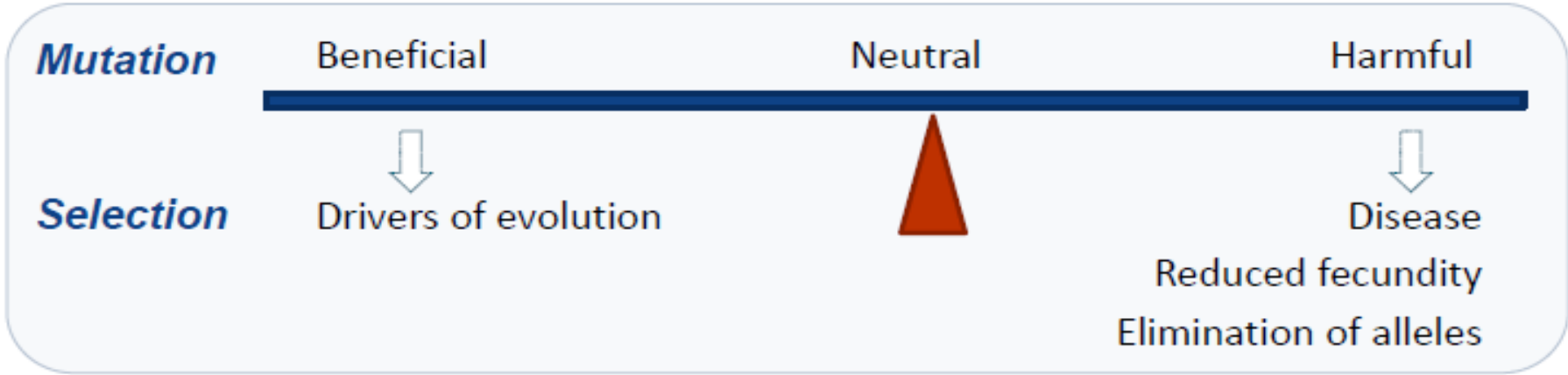
31599 structural variants (>50bp)

156 inversions (>50bp)

Origin and frequency of de novo variation

Selection-mutation balance

"balance between genetic copying errors that turn normal alleles into harmful mutations, and selection eliminating these mutations"



Frequency of de novo mutations

Estimation per generation mutation rate

7.6×10^{-9} to 2.2×10^{-8} = **50-100 *de novo* mutations per genome**

These mutations are under limited selective pressure!

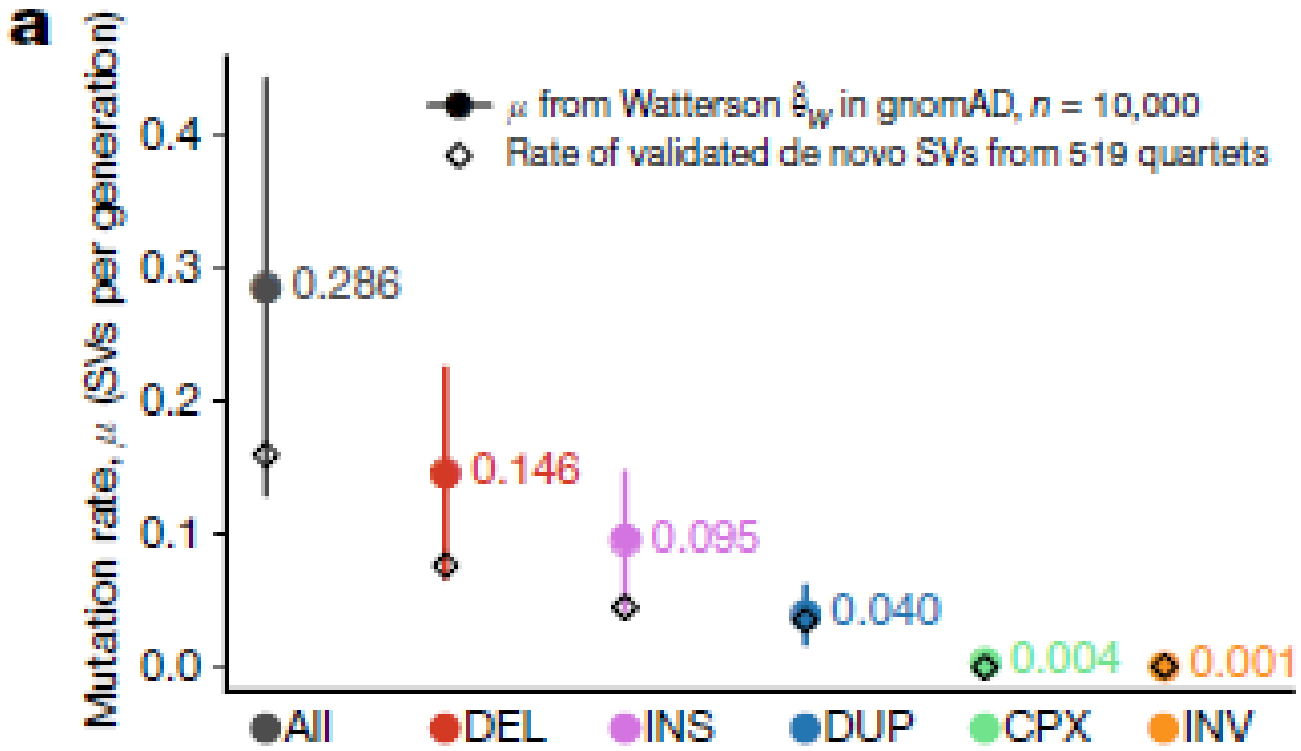
Estimated **de novo mutations per exome**:
1.4 exonic mutatons/ individual

Frequency of de novo structural variation

0.29 de novo SVs per generation in regions of the genome accessible to short-read WGS or 1 per 2-8 live births

Collins et al., Nature, 2020

Frequency varies along types of SVs



Collins et al., Nature, 2020

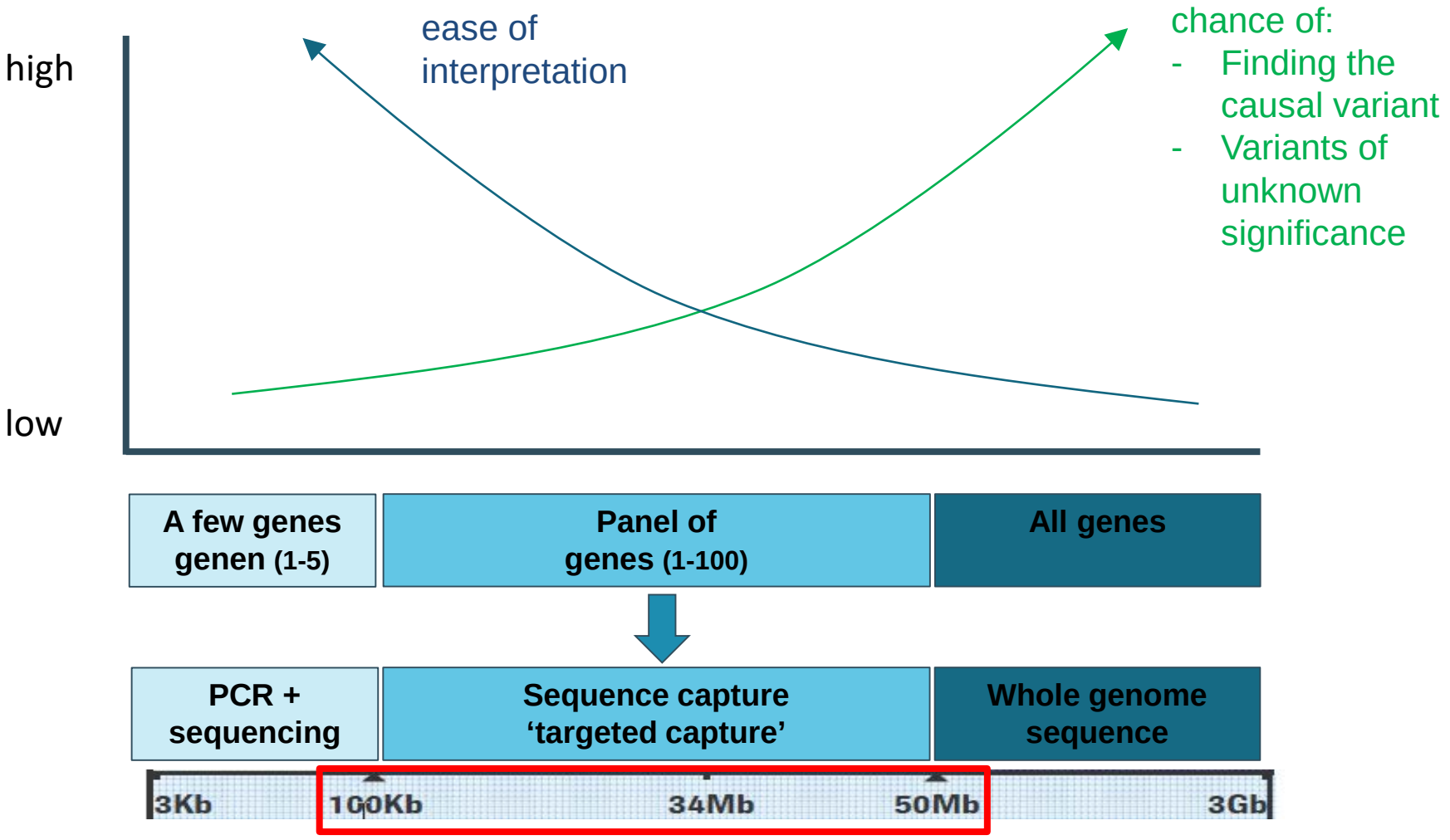
Overview

Variation Detected in a Typical Human Genome

Individuals vary greatly in a wide range of biological functions, determined in part by variation among their genomes. Any individual genome will contain the following:

- ≈5-10 million SNPs (varies by population)
- 25,000-50,000 rare variants (private mutations or seen previously in < 0.5% of individuals tested)
- ≈75 new base pair mutations not detected in parental genomes
- 3-7 new CNVs involving ≈500 kb of DNA
- 200,000-500,000 indels (1-50 bp) (varies by population)
- 500-1000 deletions 1-45 kb, overlapping ≈200 genes
- ≈150 in-frame indels
- ≈200-250 shifts in reading frame
- 10,000-12,000 synonymous SNPs
- 8,000-11,000 nonsynonymous SNPs in 4,000-5,000 genes
- 175-500 rare nonsynonymous variants
- 1 new nonsynonymous mutation
- ≈100 premature stop codons
- 40-50 splice site-disrupting variants
- 250-300 genes with likely loss-of-function variants
- ≈25 genes predicted to be completely inactivated

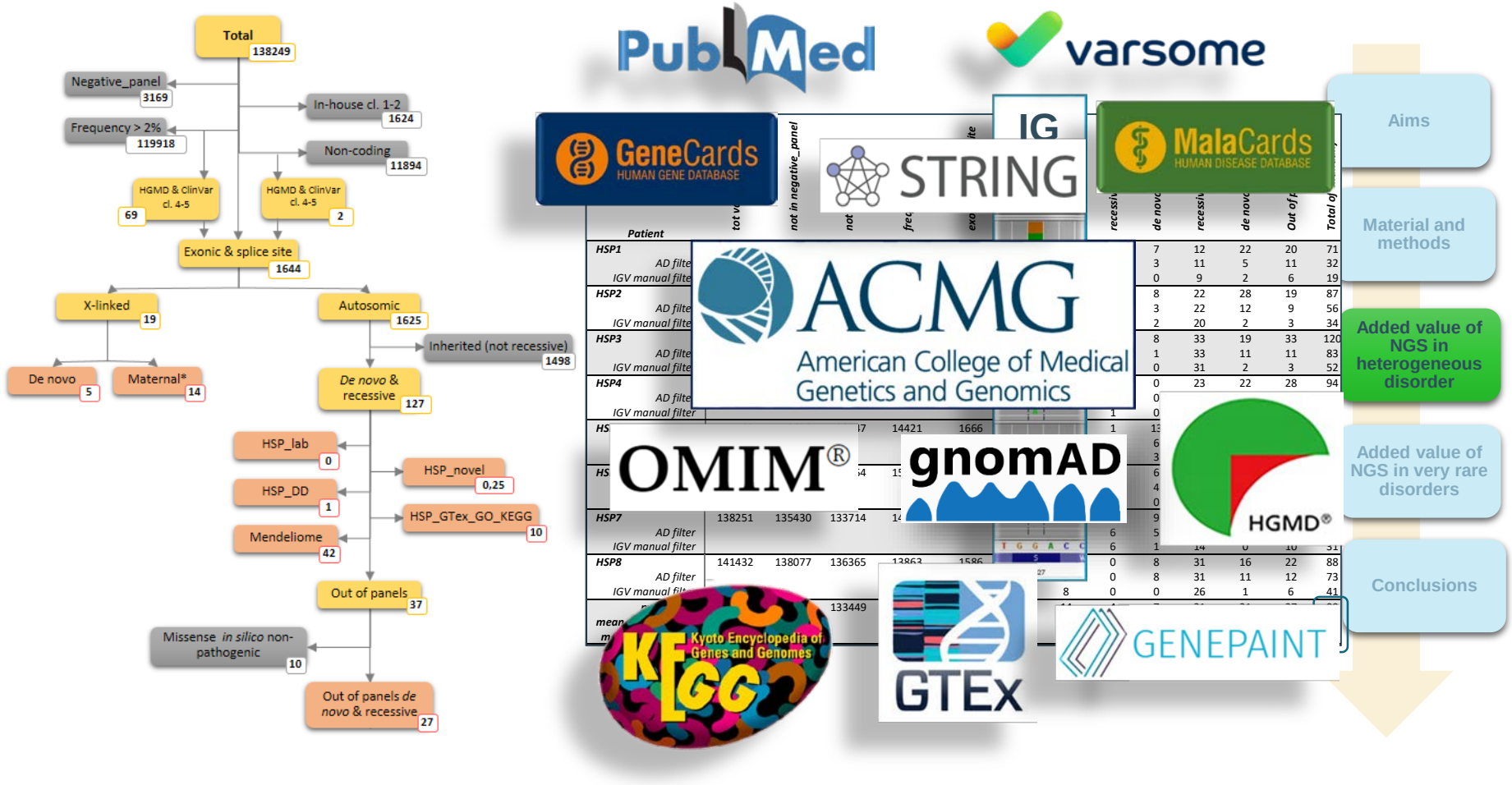
Molecular diagnostics WES/WGS



Interpretation of variants

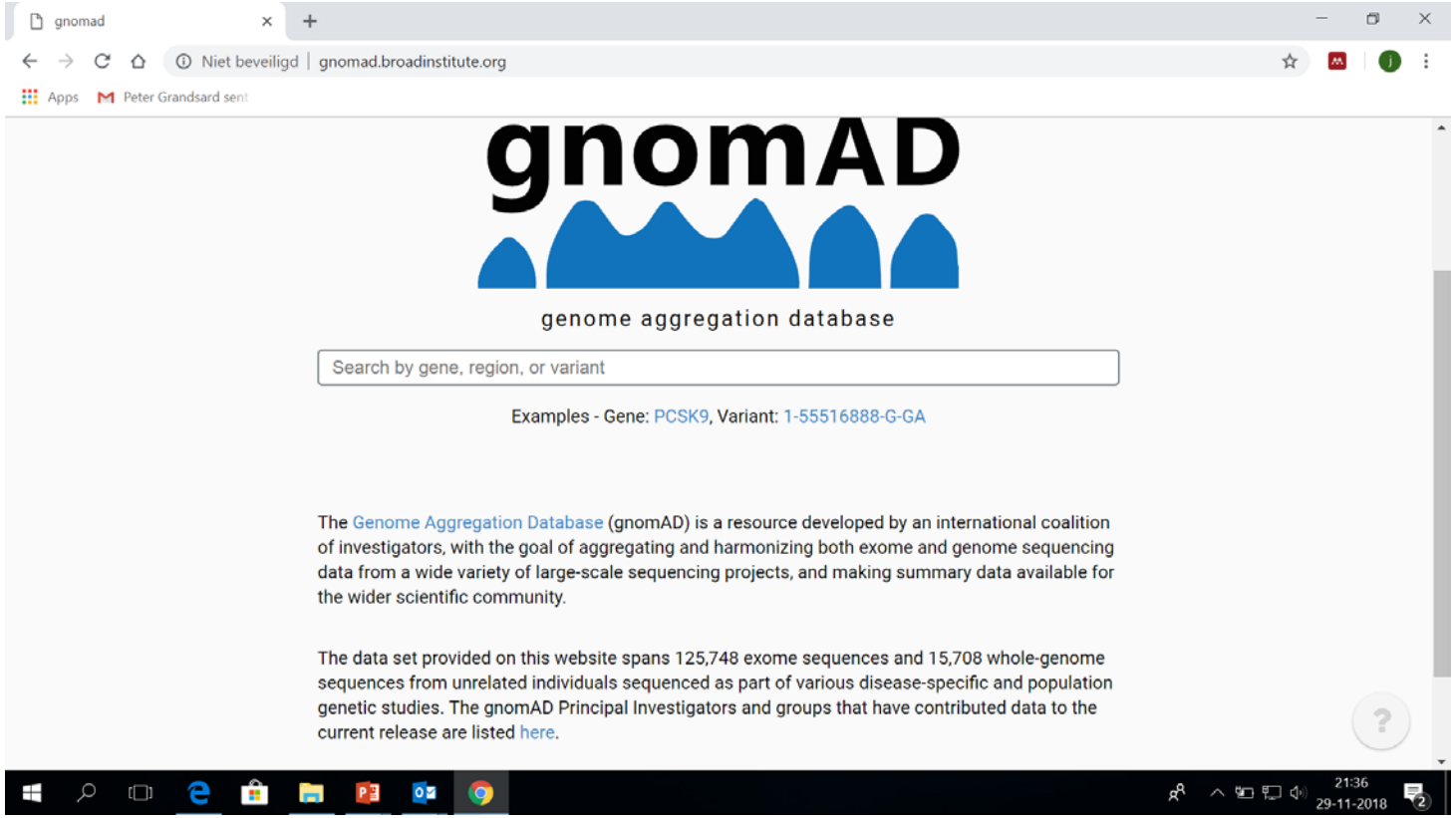
- Databases
  - SNP Databases > population frequenties
  - (internationale en lokale) mutation databases

HSP: trio-based WES analysis of a cohort of 8 patients



gnomAD (past Exac) database

Gene identification from genome wide population sequencing data based on 140000 exomes

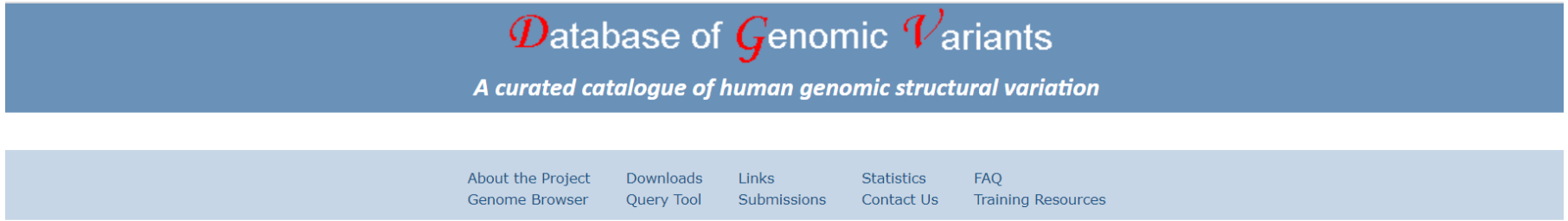


Probability of being LOF intolerant

- Haploinsufficiency to estimate the total number of autosomal recessive human protein-coding genes based on mutation tolerance
- Haploinsufficient genes do not tolerate loss-off-function (LOF) variants in one of the two alleles.
- Their probability of being LOF intolerant (pLI) is thus close to 1.
- In Gnomad pLI is measure by analysis of 140k exomes

Database of genomic variants

(curated structural variation)



Human Gene mutation database

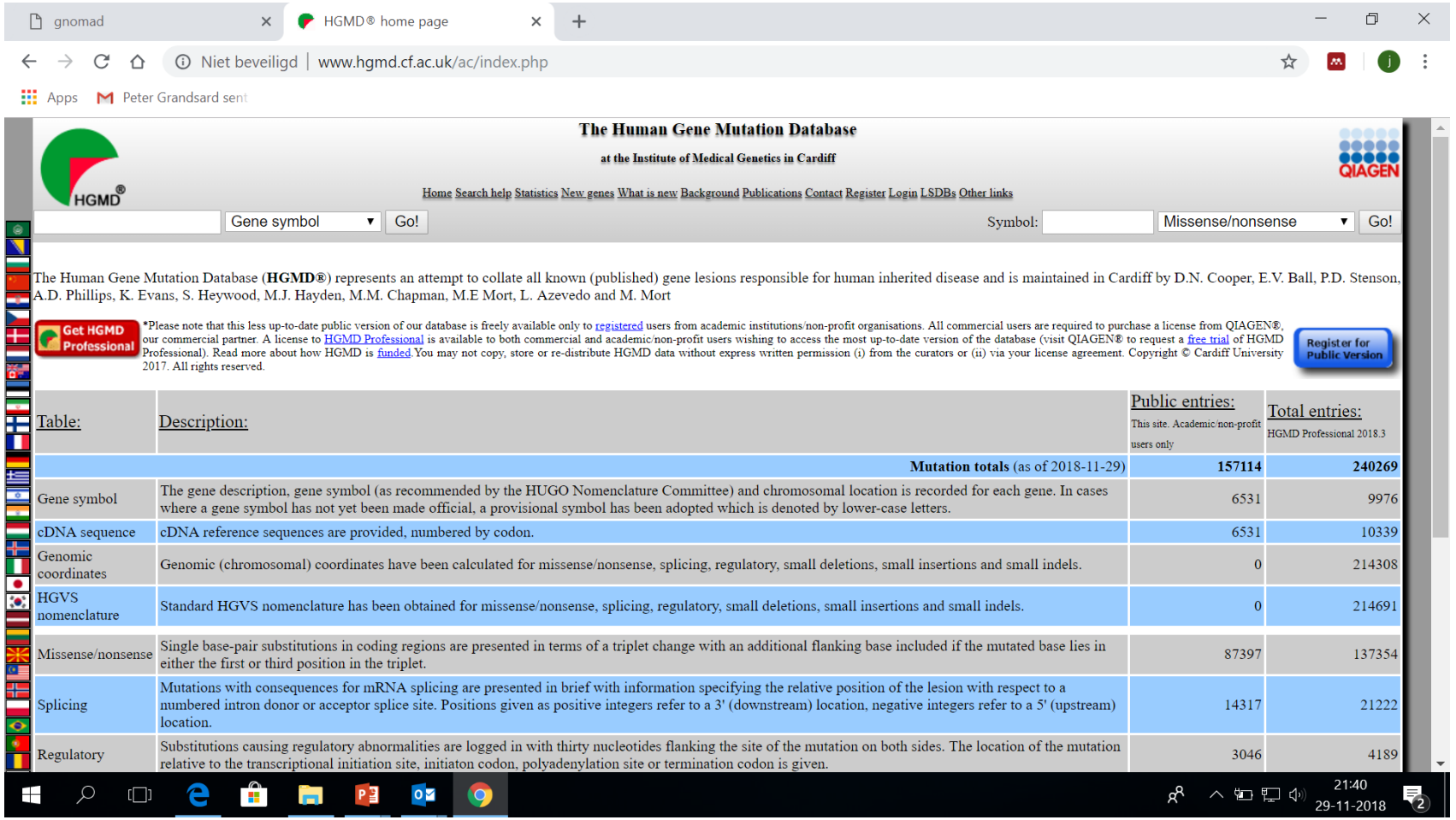
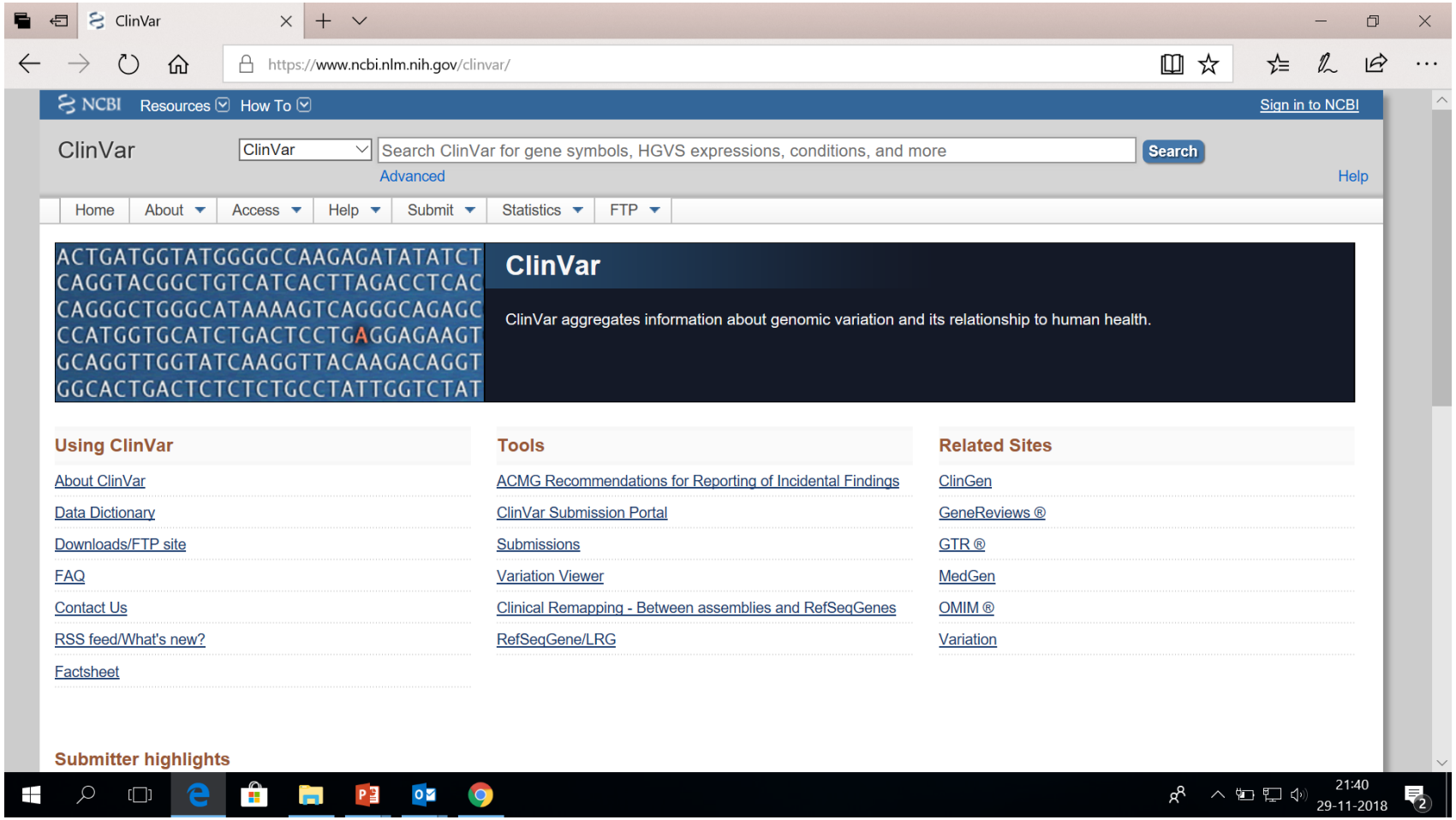


Table:	Description:	Public entries: This site. Academic/non-profit users only	Total entries: HGMD Professional 2018.3
	Mutation totals (as of 2018-11-29)	**157114**	**240269**
Gene symbol	The gene description, gene symbol (as recommended by the HUGO Nomenclature Committee) and chromosomal location is recorded for each gene. In cases where a gene symbol has not yet been made official, a provisional symbol has been adopted which is denoted by lower-case letters.	6531	9976
cDNA sequence	cDNA reference sequences are provided, numbered by codon.	6531	10339
Genomic coordinates	Genomic (chromosomal) coordinates have been calculated for missense/nonsense, splicing, regulatory, small deletions, small insertions and small indels.	0	214308
HGVS nomenclature	Standard HGVS nomenclature has been obtained for missense/nonsense, splicing, regulatory, small deletions, small insertions and small indels.	0	214691
Missense/nonsense	Single base-pair substitutions in coding regions are presented in terms of a triplet change with an additional flanking base included if the mutated base lies in either the first or third position in the triplet.	87397	137354
Splicing	Mutations with consequences for mRNA splicing are presented in brief with information specifying the relative position of the lesion with respect to a numbered intron donor or acceptor splice site. Positions given as positive integers refer to a 3' (downstream) location, negative integers refer to a 5' (upstream) location.	14317	21222
Regulatory	Substitutions causing regulatory abnormalities are logged in with thirty nucleotides flanking the site of the mutation on both sides. The location of the mutation relative to the transcriptional initiation site, initiaton codon, polyadenylation site or termination codon is given.	3046	4189

Clinvar


ClinVar
https://www.ncbi.nlm.nih.gov/clinvar/
NCBI Resources How To
Sign in to NCBI
ClinVar
ClinVar
Search ClinVar for gene symbols, HGVS expressions, conditions, and more
Search
Advanced
Help
Home About Access Help Submit Statistics FTP
ACTGATGGTATGGGGCCAAGAGATATATCT
CAGGTACGGCTGTCATCACTTAGACCTCAC
CAGGGCTGGGCATAAAAGTCAGGGCAGAGC
CCATGGTGCATCTGACTCCTGAGGAGAAGT
GCAGGTTGGTATCAAGGTTACAAGACAGGT
GGCACTGACTCTCTCTGCCTATTGGTCTAT
ClinVar
ClinVar aggregates information about genomic variation and its relationship to human health.
Using ClinVar
About ClinVar
Data Dictionary
Downloads/FTP site
FAQ
Contact Us
RSS feed/What's new?
Factsheet
Tools
ACMG Recommendations for Reporting of Incidental Findings
ClinVar Submission Portal
Submissions
Variation Viewer
Clinical Remapping - Between assemblies and RefSeqGenes
RefSeqGene/LRG
Related Sites
ClinGen
GeneReviews ®
GTR ®
MedGen
OMIM ®
Variation
Submitter highlights
21:40
29-11-2018

HUMAN MUTATION Database in BRIEF 31: E1801-E1810 (2010) Online

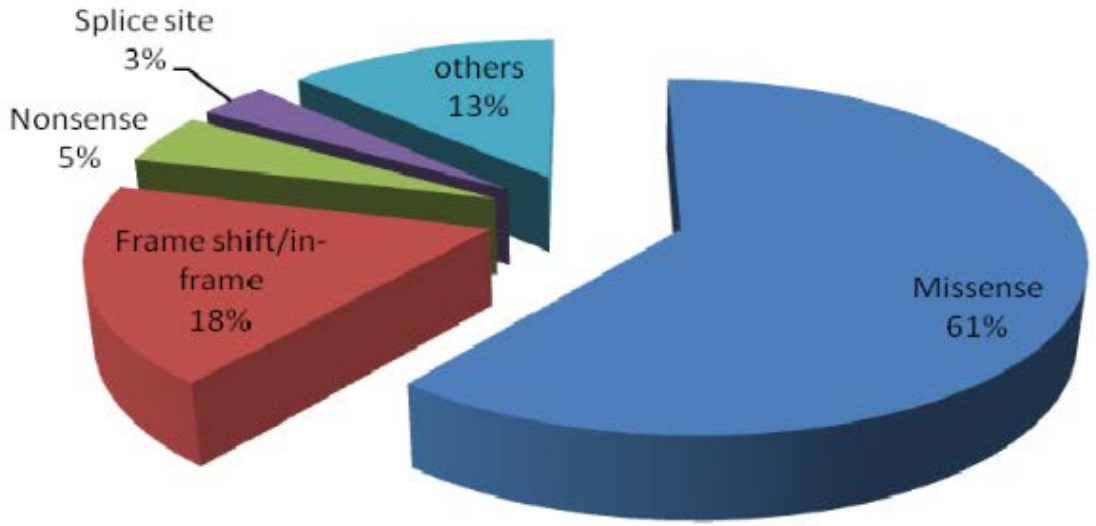
DATABASE IN BRIEF

HUMAN MUTATION

OFFICIAL JOURNAL
HGVS
HUMAN GENOME VARIATION SOCIETY
www.hgvs.org

LQTS Gene LOVD Database

Tao Zhang[1,2*], Arthur Moss[3,*], Peikuan Cong[2,*], Min Pan [2,*], Bingxi Chang[4], Liangrong Zheng[5], Quan Fang[4], Wojciech Zareba[3], Jennifer Robinson[3], Changsong Lin[2], Zhongxiang Li[6], Junfang Wei[7], Qiang Zeng[8], Long QT International Registry Investigators, HVP-China Investigators, and Ming Qi[1,2,9**]



ACMG STANDARDS AND GUIDELINES | Genetics in Medicine

Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology

Sue Richards, PhD[1], Nazneen Aziz, PhD[2,16], Sherri Bale, PhD[3], David Bick, MD[4], Soma Das, PhD[5], Julie Gastier-Foster, PhD[6,7,8], Wayne W. Grody, MD, PhD[9,10,11], Madhuri Hegde, PhD[12], Elaine Lyon, PhD[13], Elaine Spector, PhD[14], Karl Voelkerding, MD[13] and Heidi L. Rehm, PhD[15]; on behalf of the ACMG Laboratory Quality Assurance Committee

Class of risk	Clinical significance
1	not patogenic
2	likely not pathogenic
3	uncertain
4	likely pathogenic
5	definitely patogenic

Plon SE, Eccles DM, Easton D, et al. Sequence variant classification and reporting: recommendations for improving the interpretation of cancer susceptibility genetic test results. Hum Mutat. 2008; 29:1282–1291. [PubMed: 18951446]

ACMG STANDARDS AND GUIDELINES

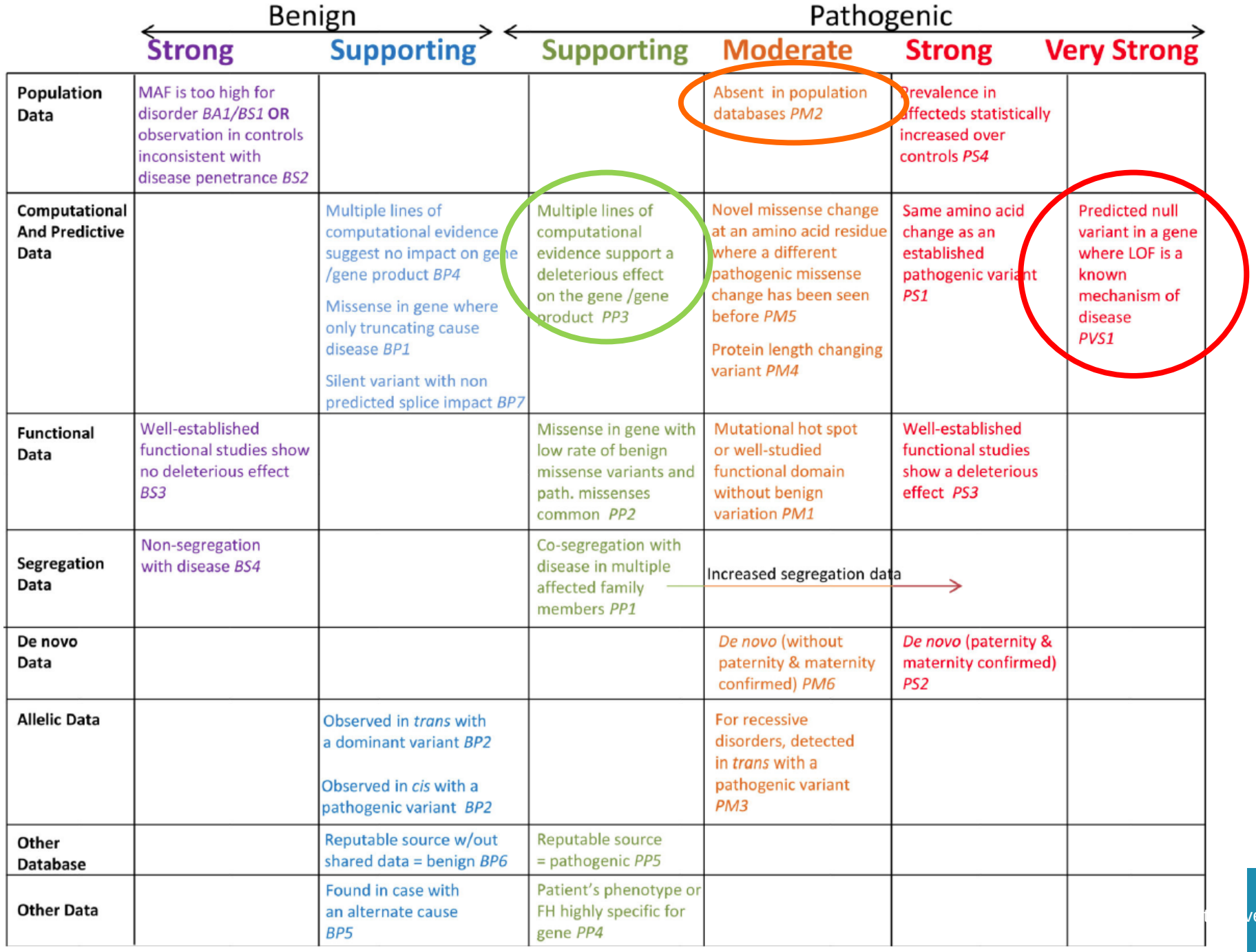
Table 3 Criteria for classifying pathogenic variants

Evidence of pathogenicity	Category
Very strong	PVS1 null variant (nonsense, frameshift, canonical ±1 or 2 splice sites, initiation codon, single or multiexon deletion) in a gene where LOF is a known mechanism of disease Caveats: • Beware of genes where LOF is not a known disease mechanism (e.g., *GFAP*, *MYH7*) • Use caution interpreting LOF variants at the extreme 3′ end of a gene • Use caution with splice variants that are predicted to lead to exon skipping but leave the remainder of the protein intact • Use caution in the presence of multiple transcripts
Strong	PS1 Same amino acid change as a previously established pathogenic variant regardless of nucleotide change Example: Val→Leu caused by either G>C or G>T in the same codon Caveat: Beware of changes that impact splicing rather than at the amino acid/protein level PS2 De novo (both maternity and paternity confirmed) in a patient with the disease and no family history Note: Confirmation of paternity only is insufficient. Egg donation, surrogate motherhood, errors in embryo transfer, and so on, can contribute to nonmaternity. PS3 Well-established in vitro or in vivo functional studies supportive of a damaging effect on the gene or gene product Note: Functional studies that have been validated and shown to be reproducible and robust in a clinical diagnostic laboratory setting are considered the most well established. PS4 The prevalence of the variant in affected individuals is significantly increased compared with the prevalence in controls Note 1: Relative risk or OR, as obtained from case–control studies, is >5.0, and the confidence interval around the estimate of relative risk or OR does not include 1.0. See the article for detailed guidance. Note 2: In instances of very rare variants where case–control studies may not reach statistical significance, the prior observation of the variant in multiple unrelated patients with the same phenotype, and its absence in controls, may be used as moderate level of evidence.
Moderate	PM1 Located in a mutational hot spot and/or critical and well-established functional domain (e.g., active site of an enzyme) without benign variation

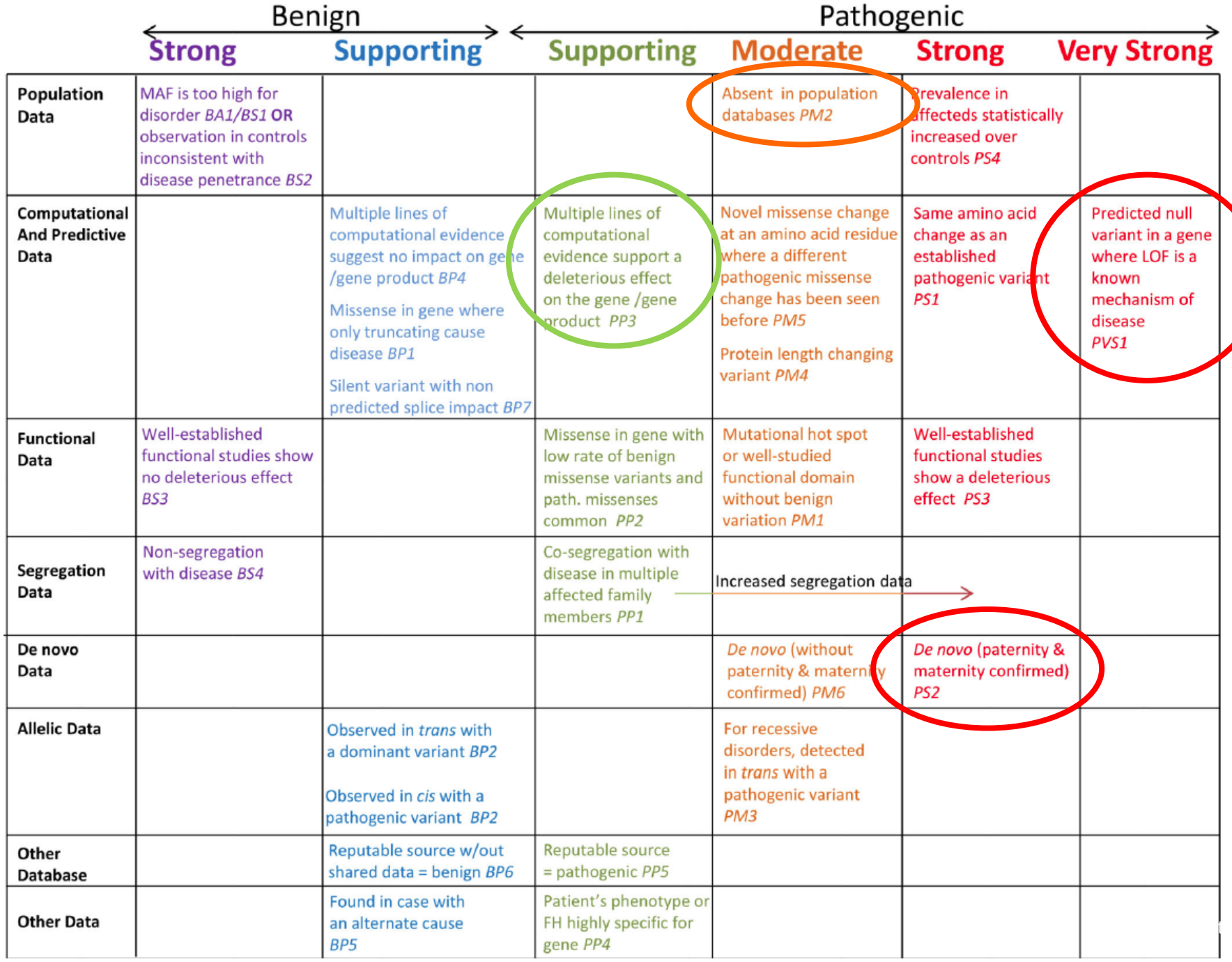
	Benign		Pathogenic			
	Strong	**Supporting**	**Supporting**	**Moderate**	**Strong**	**Very Strong**
Population Data	MAF is too high for disorder *BA1/BS1* **OR** observation in controls inconsistent with disease penetrance *BS2*			Absent in population databases *PM2*	Prevalence in affecteds statistically increased over controls *PS4*	
Computational And Predictive Data		Multiple lines of computational evidence suggest no impact on gene /gene product *BP4* Missense in gene where only truncating cause disease *BP1* Silent variant with non predicted splice impact *BP7*	Multiple lines of computational evidence support a deleterious effect on the gene /gene product *PP3*	Novel missense change at an amino acid residue where a different pathogenic missense change has been seen before *PM5* Protein length changing variant *PM4*	Same amino acid change as an established pathogenic variant *PS1*	Predicted null variant in a gene where LOF is a known mechanism of disease *PVS1*
Functional Data	Well-established functional studies show no deleterious effect *BS3*		Missense in gene with low rate of benign missense variants and path. missenses common *PP2*	Mutational hot spot or well-studied functional domain without benign variation *PM1*	Well-established functional studies show a deleterious effect *PS3*	
Segregation Data	Non-segregation with disease *BS4*		Co-segregation with disease in multiple affected family members *PP1*	Increased segregation data →		
De novo Data				*De novo* (without paternity & maternity confirmed) *PM6*	*De novo* (paternity & maternity confirmed) *PS2*	
Allelic Data		Observed in *trans* with a dominant variant *BP2* Observed in *cis* with a pathogenic variant *BP2*		For recessive disorders, detected in *trans* with a pathogenic variant *PM3*		
Other Database		Reputable source w/out shared data = benign *BP6*	Reputable source = pathogenic *PP5*			
Other Data		Found in case with an alternate cause *BP5*	Patient's phenotype or FH highly specific for gene *PP4*			

	Benign		Pathogenic			
	Strong	**Supporting**	**Supporting**	**Moderate**	**Strong**	**Very Strong**
Population Data	MAF is too high for disorder *BA1/BS1* **OR** observation in controls inconsistent with disease penetrance *BS2*			Absent in population databases *PM2*	Prevalence in affecteds statistically increased over controls *PS4*	
Computational And Predictive Data		Multiple lines of computational evidence suggest no impact on gene /gene product *BP4* Missense in gene where only truncating cause disease *BP1* Silent variant with non predicted splice impact *BP7*	Multiple lines of computational evidence support a deleterious effect on the gene /gene product *PP3*	Novel missense change at an amino acid residue where a different pathogenic missense change has been seen before *PM5* Protein length changing variant *PM4*	Same amino acid change as an established pathogenic variant *PS1*	Predicted null variant in a gene where LOF is a known mechanism of disease *PVS1*
Functional Data	Well-established functional studies show no deleterious effect *BS3*		Missense in gene with low rate of benign missense variants and path. missenses common *PP2*	Mutational hot spot or well-studied functional domain without benign variation *PM1*	Well-established functional studies show a deleterious effect *PS3*	
Segregation Data	Non-segregation with disease *BS4*		Co-segregation with disease in multiple affected family members *PP1*	Increased segregation data →		
De novo Data				*De novo* (without paternity & maternity confirmed) *PM6*	*De novo* (paternity & maternity confirmed) *PS2*	
Allelic Data		Observed in *trans* with a dominant variant *BP2* Observed in *cis* with a pathogenic variant *BP2*		For recessive disorders, detected in *trans* with a pathogenic variant *PM3*		
Other Database		Reputable source w/out shared data = benign *BP6*	Reputable source = pathogenic *PP5*			
Other Data		Found in case with an alternate cause *BP5*	Patient's phenotype or FH highly specific for gene *PP4*			

	Benign		Pathogenic			
	Strong	**Supporting**	**Supporting**	**Moderate**	**Strong**	**Very Strong**
Population Data	MAF is too high for disorder *BA1/BS1* **OR** observation in controls inconsistent with disease penetrance *BS2*			Absent in population databases *PM2*	Prevalence in affecteds statistically increased over controls *PS4*	
Computational And Predictive Data		Multiple lines of computational evidence suggest no impact on gene /gene product *BP4* Missense in gene where only truncating cause disease *BP1* Silent variant with non predicted splice impact *BP7*	Multiple lines of computational evidence support a deleterious effect on the gene /gene product *PP3*	Novel missense change at an amino acid residue where a different pathogenic missense change has been seen before *PM5* Protein length changing variant *PM4*	Same amino acid change as an established pathogenic variant *PS1*	Predicted null variant in a gene where LOF is a known mechanism of disease *PVS1*
Functional Data	Well-established functional studies show no deleterious effect *BS3*		Missense in gene with low rate of benign missense variants and path. missenses common *PP2*	Mutational hot spot or well-studied functional domain without benign variation *PM1*	Well-established functional studies show a deleterious effect *PS3*	
Segregation Data	Non-segregation with disease *BS4*		Co-segregation with disease in multiple affected family members *PP1*	Increased segregation data →		
De novo Data				*De novo* (without paternity & maternity confirmed) *PM6*	*De novo* (paternity & maternity confirmed) *PS2*	
Allelic Data		Observed in *trans* with a dominant variant *BP2* Observed in *cis* with a pathogenic variant *BP2*		For recessive disorders, detected in *trans* with a pathogenic variant *PM3*		
Other Database		Reputable source w/out shared data = benign *BP6*	Reputable source = pathogenic *PP5*			
Other Data		Found in case with an alternate cause *BP5*	Patient's phenotype or FH highly specific for gene *PP4*			



	Benign		Pathogenic			
	Strong	**Supporting**	**Supporting**	**Moderate**	**Strong**	**Very Strong**
Population Data	MAF is too high for disorder *BA1/BS1* **OR** observation in controls inconsistent with disease penetrance *BS2*			Absent in population databases *PM2*	Prevalence in affecteds statistically increased over controls *PS4*	
Computational And Predictive Data		Multiple lines of computational evidence suggest no impact on gene /gene product *BP4* Missense in gene where only truncating cause disease *BP1* Silent variant with non predicted splice impact *BP7*	Multiple lines of computational evidence support a deleterious effect on the gene /gene product *PP3*	Novel missense change at an amino acid residue where a different pathogenic missense change has been seen before *PM5* Protein length changing variant *PM4*	Same amino acid change as an established pathogenic variant *PS1*	Predicted null variant in a gene where LOF is a known mechanism of disease *PVS1*
Functional Data	Well-established functional studies show no deleterious effect *BS3*		Missense in gene with low rate of benign missense variants and path. missenses common *PP2*	Mutational hot spot or well-studied functional domain without benign variation *PM1*	Well-established functional studies show a deleterious effect *PS3*	
Segregation Data	Non-segregation with disease *BS4*		Co-segregation with disease in multiple affected family members *PP1*	Increased segregation data →		
De novo Data				*De novo* (without paternity & maternity confirmed) *PM6*	*De novo* (paternity & maternity confirmed) *PS2*	
Allelic Data		Observed in *trans* with a dominant variant *BP2* Observed in *cis* with a pathogenic variant *BP2*		For recessive disorders, detected in *trans* with a pathogenic variant *PM3*		
Other Database		Reputable source w/out shared data = benign *BP6*	Reputable source = pathogenic *PP5*			
Other Data		Found in case with an alternate cause *BP5*	Patient's phenotype or FH highly specific for gene *PP4*			



	Benign		Pathogenic			
	Strong	**Supporting**	**Supporting**	**Moderate**	**Strong**	**Very Strong**
Population Data	MAF is too high for disorder *BA1/BS1* **OR** observation in controls inconsistent with disease penetrance *BS2*			Absent in population databases *PM2*	Prevalence in affecteds statistically increased over controls *PS4*	
Computational And Predictive Data		Multiple lines of computational evidence suggest no impact on gene /gene product *BP4* Missense in gene where only truncating cause disease *BP1* Silent variant with non predicted splice impact *BP7*	Multiple lines of computational evidence support a deleterious effect on the gene /gene product *PP3*	Novel missense change at an amino acid residue where a different pathogenic missense change has been seen before *PM5* Protein length changing variant *PM4*	Same amino acid change as an established pathogenic variant *PS1*	Predicted null variant in a gene where LOF is a known mechanism of disease *PVS1*
Functional Data	Well-established functional studies show no deleterious effect *BS3*		Missense in gene with low rate of benign missense variants and path. missenses common *PP2*	Mutational hot spot or well-studied functional domain without benign variation *PM1*	Well-established functional studies show a deleterious effect *PS3*	
Segregation Data	Non-segregation with disease *BS4*		Co-segregation with disease in multiple affected family members *PP1*	Increased segregation data →		
De novo Data				*De novo* (without paternity & maternity confirmed) *PM6*	*De novo* (paternity & maternity confirmed) *PS2*	
Allelic Data		Observed in *trans* with a dominant variant *BP2* Observed in *cis* with a pathogenic variant *BP2*		For recessive disorders, detected in *trans* with a pathogenic variant *PM3*		
Other Database		Reputable source w/out shared data = benign *BP6*	Reputable source = pathogenic *PP5*			
Other Data		Found in case with an alternate cause *BP5*	Patient's phenotype or FH highly specific for gene *PP4*			

Table 5 Rules for combining criteria to classify sequence variants

Pathogenic	(i) 1 Very strong (PVS1) *AND* (a) ≥1 Strong (PS1–PS4) *OR* (b) ≥2 Moderate (PM1–PM6) *OR* (c) 1 Moderate (PM1–PM6) and 1 supporting (PP1–PP5) *OR* (d) ≥2 Supporting (PP1–PP5) (ii) ≥2 Strong (PS1–PS4) *OR* (iii) 1 Strong (PS1–PS4) *AND* (a)≥3 Moderate (PM1–PM6) *OR* (b)2 Moderate (PM1–PM6) *AND* ≥2 Supporting (PP1–PP5) *OR* (c)1 Moderate (PM1–PM6) *AND* ≥4 supporting (PP1–PP5)
Likely pathogenic	(i) 1 Very strong (PVS1) *AND* 1 moderate (PM1–PM6) *OR* (ii) 1 Strong (PS1–PS4) *AND* 1–2 moderate (PM1–PM6) *OR* (iii) 1 Strong (PS1–PS4) *AND* ≥2 supporting (PP1–PP5) *OR* (iv) ≥3 Moderate (PM1–PM6) *OR* (v) 2 Moderate (PM1–PM6) *AND* ≥2 supporting (PP1–PP5) *OR* (vi) 1 Moderate (PM1–PM6) *AND* ≥4 supporting (PP1–PP5)
Benign	(i) 1 Stand-alone (BA1) *OR* (ii) ≥2 Strong (BS1–BS4)
Likely benign	(i) 1 Strong (BS1–BS4) and 1 supporting (BP1–BP7) *OR* (ii) ≥2 Supporting (BP1–BP7)
Uncertain significance	(i) Other criteria shown above are not met OR (ii) the criteria for benign and pathogenic are contradictory

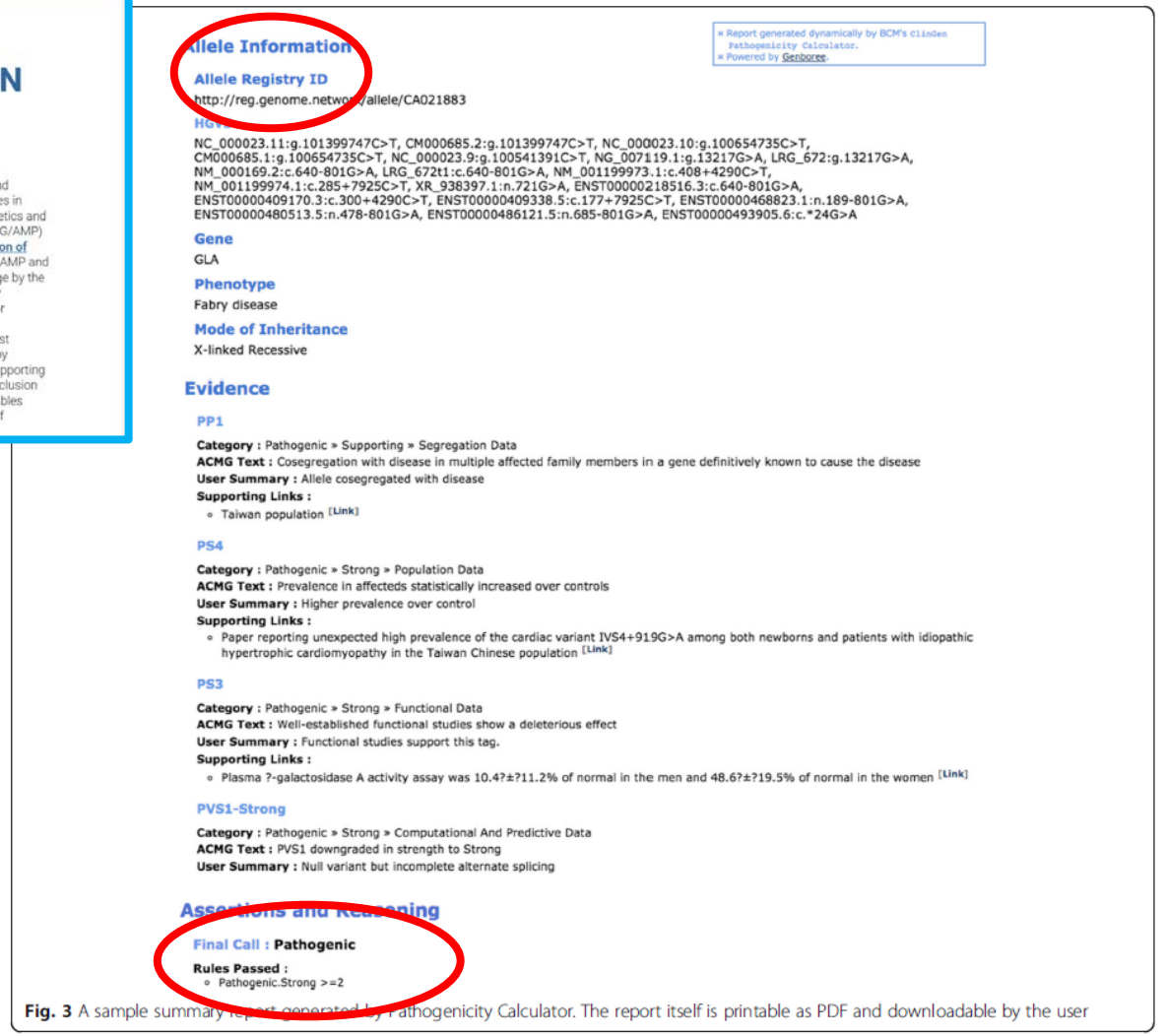


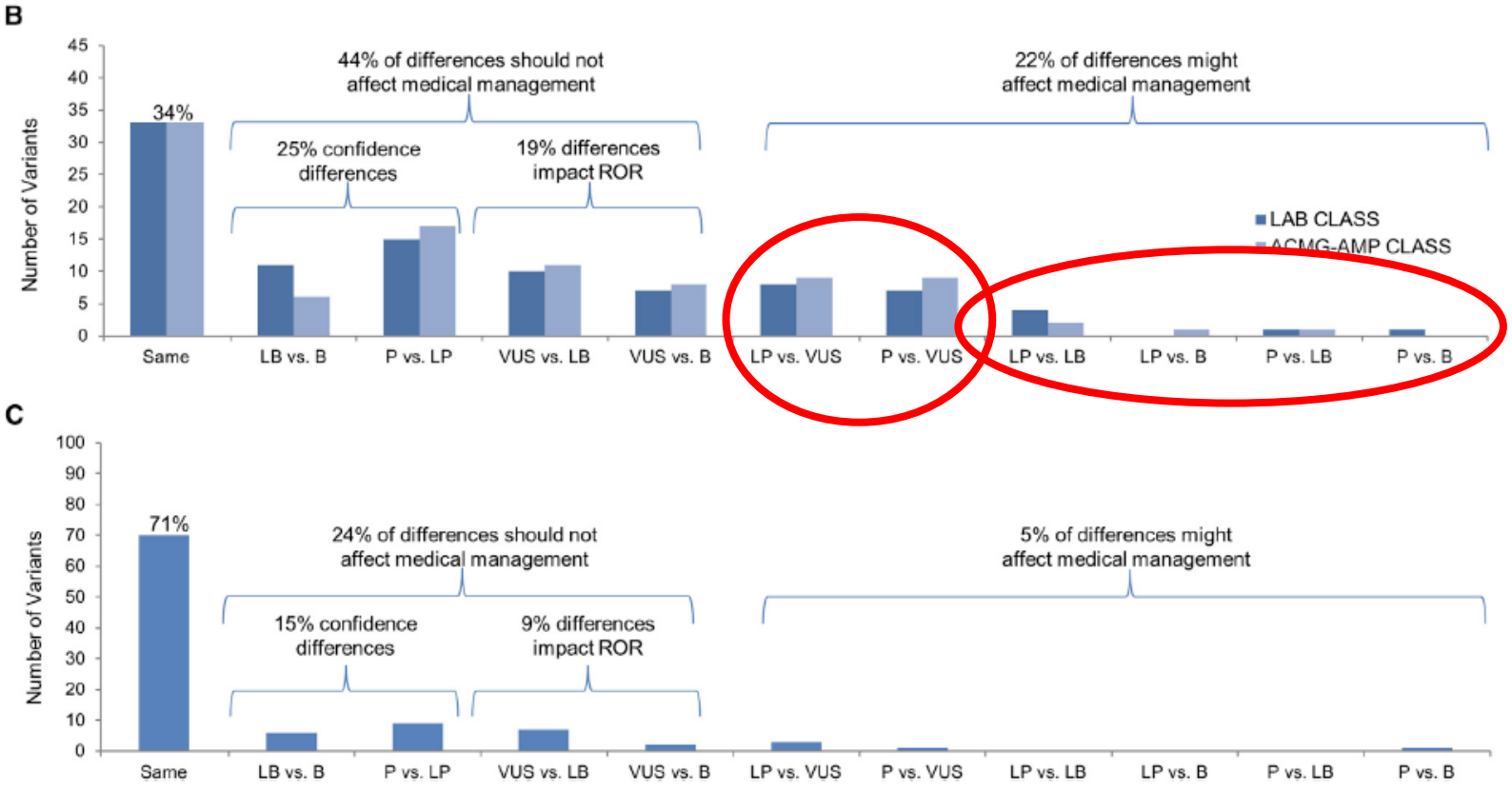
Fig. 3 A sample summary report generated by Pathogenicity Calculator. The report itself is printable as PDF and downloadable by the user

Patel et al. Genome Medicine 2017

ARTICLE

Performance of ACMG-AMP Variant-Interpretation Guidelines among Nine Laboratories in the Clinical Sequencing Exploratory Research Consortium

Laura M. Amendola,[1,16] Gail P. Jarvik,[1,16,*] Michael C. Leo,[2] Heather M. McLaughlin,[3] Yassmine Akkari,[4] Michelle D. Amaral,[5] Jonathan S. Berg,[6] Sawona Biswas,[7] Kevin M. Bowling,[5] Laura K. Conlin,[7] Greg M. Cooper,[5] Michael O. Dorschner,[8] Matthew C. Dulik,[9] Arezou A. Ghazani,[10] Rajarshi Ghosh,[11] Robert C. Green,[3,12,15] Ragan Hart,[1] Carrie Horton,[13] Jennifer J. Johnston,[14] Matthew S. Lebo,[3,12] Aleksandar Milosavljevic,[11] Jeffrey Ou,[1] Christine M. Pak,[4] Ronak Y. Patel,[11] Sumit Punj,[4] Carolyn Sue Richards,[4] Joseph Salama,[1] Natasha T. Strande,[6] Yaping Yang,[11] Sharon E. Plon,[11] Leslie G. Biesecker,[14] and Heidi L. Rehm[3,12,15,*]



(B) Inter-laboratory concordance of 97 variants. This graph compares the same calls, based on either the ACMG-AMP rules or the site's rules, between laboratories.
(C) Inter-laboratory concordance after consensus efforts. This graph shows a final comparison of calls between sites after consensus-building efforts.

Table 1. ACMG-AMP Rule Clarifications and Suggestions for Modification

Rule	Description	Clarifications and/or Suggestions
PVS1	variant predicted null where LOF is a mechanism of disease	do not apply to variants that are near the 3′ end of the gene and escape nonsense-mediated decay
PS1	variant with the same amino acid change as a previously established pathogenic variant, regardless of nucleotide change	does not include the same variant being assessed because it is not yet pathogenic, and the rule is intended for variants with a different nucleotide change
PS2	de novo variant with confirmed maternity and paternity	apply this rule as moderate or supporting if the variant is mosaic and its frequency in tissue is consistent with the phenotype
PS3	variant shown to have a deleterious effect by a well-established functional study	reduce the strength for assays that are not as well validated or linked to the phenotype
PM1	variant located in a mutational hotspot and/or critical and well-established functional domain	not meant for truncations; more clarification is needed for applying this rule
PM2, BS1	variant absent in population databases or with an allele frequency too high for the disease	cannot assume longer indels would be detected by next-generation sequencing
		use a published control dataset if its size is at least 1,000 individuals
		cannot be applied for low-quality calls or non-covered regions
		must define the condition and inheritance pattern
PM3	for recessive disorders, variant in *trans* with a pathogenic variant	invoke this rule as supporting if the phase is not established
		can upgrade if more than one proband is reported
PM4	protein-length-changing variant	applicable for in-frame deletions, insertions, or stop-loss variants, but not frameshifts, nonsense, and splice variants
PM5	novel missense variant at amino acid with different pathogenic missense change	ensure pathogenicity of previously reported variant
		suggest changing "novel" to "different" because some variants that are not novel might require assessment with this rule
PP3, BP4	variant with multiple lines of computational evidence	all lines must agree
PP4	the patient's phenotype or family history is highly specific to the genotype	not meant to be used for genetically heterogeneous conditions or conditions with unsolved etiology
		not typically applied for an analysis of incidental findings, but it could be applied for prior observations
PP5, BP6	variant called pathogenic or benign by a reputable source	only applicable when evidence is not available (e.g., Sharing Clinical Reports Project)
BS2	variant observed in a healthy adult for a disorder with full penetrance at an early age	populations might not have been screened or excluded for the phenotype
BP1	variant in a gene in which truncations primarily cause disease	clarify the meaning of "primary"; suggest >90%
BP2, BP5	variant in *trans* with a dominant pathogenic variant (BP2) or in an individual with an alternate molecular basis for disease (BP5)	clarify that one should apply BP2 when the pathogenic variant is seen in the same gene as the variant being evaluated and apply BP5 when the pathogenic variant is in a different gene

Genetics inMedicine | ORIGINAL RESEARCH ARTICLE

Clinical laboratories collaborate to resolve differences in variant interpretations submitted to ClinVar

Steven M. Harrison, PhD[1,2], Jill S. Dolinsky, MS[3], Amy E. Knight Johnson, MS[4], Tina Pesaran, MA, MS[3], Danielle R. Azzariti, MS[1], Sherri Bale, PhD[5], Elizabeth C. Chao, MD[3,6], Soma Das, PhD[4], Lisa Vincent, PhD[5] and Heidi L. Rehm, PhD[1,2,7,8]

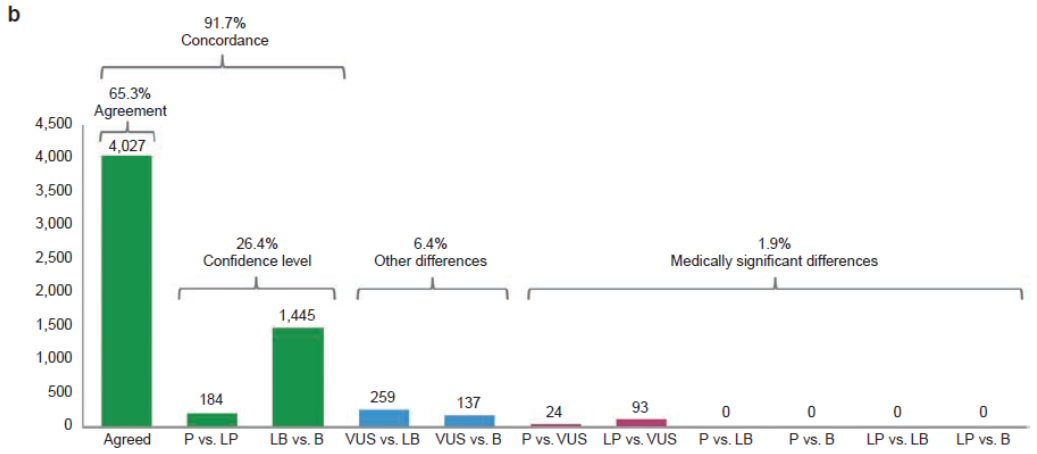


Figure 1 Distribution of variant interpretation differences between four clinical laboratories. (**a**) Interpretation comparison of data in ClinVar (as of January 1, 2016) before resolution efforts. (**b**) Interpretation comparison after reassessing 33% (242/724) of shared variants with interpretation differences.

Genetic testing

Molecular testing

1 disease
1 (single) mutation

1 disease
1 gene

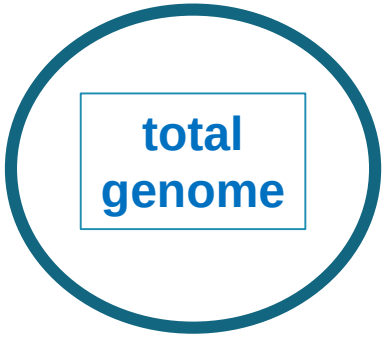
different diseases
1 (single) gene

1 disease
few genes

different diseases
many genes

1 disease
1 chromosomal anomaly

different diseases
different anomalies

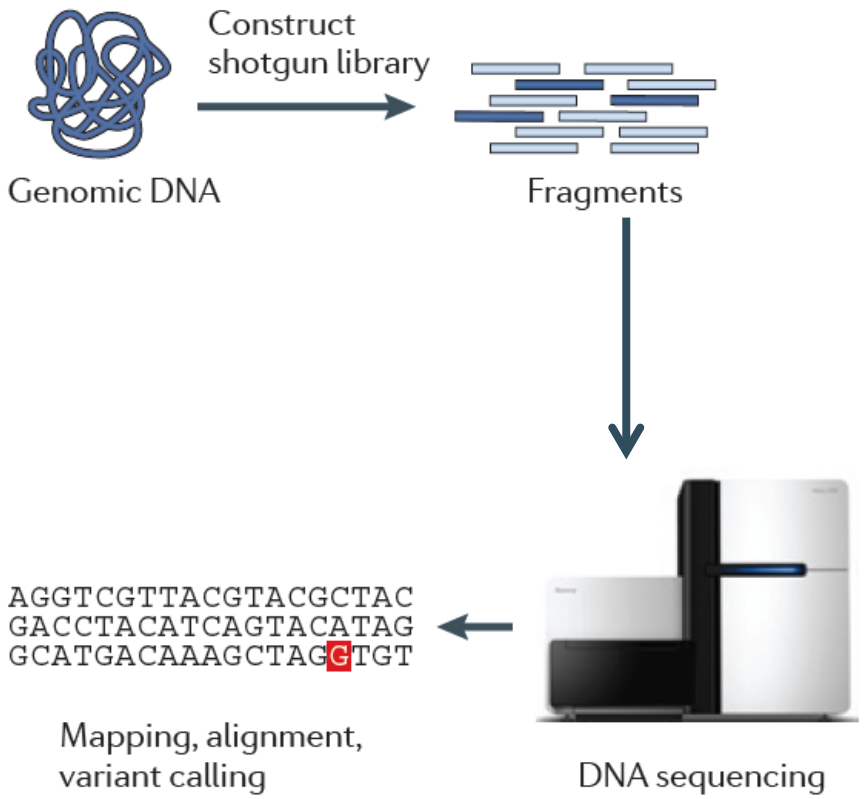


(un)known defect
karyotyping

Molecular cytogenetic testing

Cytogenetics

Whole-Genome SEQUENCING



MCA/ID: trio-based WGS analysis of 3 patients with unknown clinical diagnosis

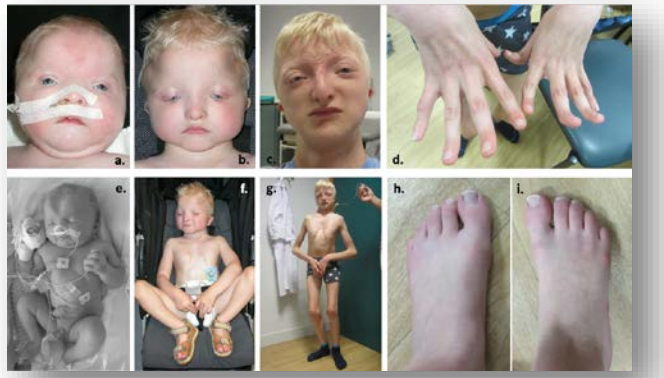
Anamnesis:

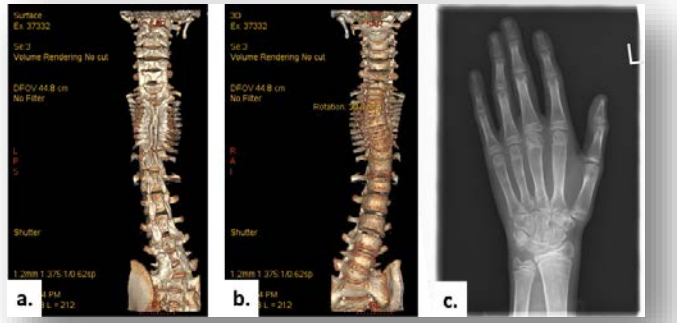
- Only child of an healthy unrelated couple.
- At birth facial dysmorphism with polymalformative syndrome
- Large anterior fontanella
- Microcephaly
- Right-turning
- Abdominal wa
- Cleft palate
- Hypoplasia of
- Cryptorchidia
- Major hearing
- Hypertension
- Feeding problems
- Short stature
- Congenital thoracic vertebral fusion → severe torsional scoliosis

Molecular analysis:

- Karyotyping
- FISH for 22q11.2 and 9p-
- Array-CGH
  - 8q12.1(56899737-57048789)x3mat
  - 16p13.3(4379999-4443009)x3mat
  - likely benign

Patient 1





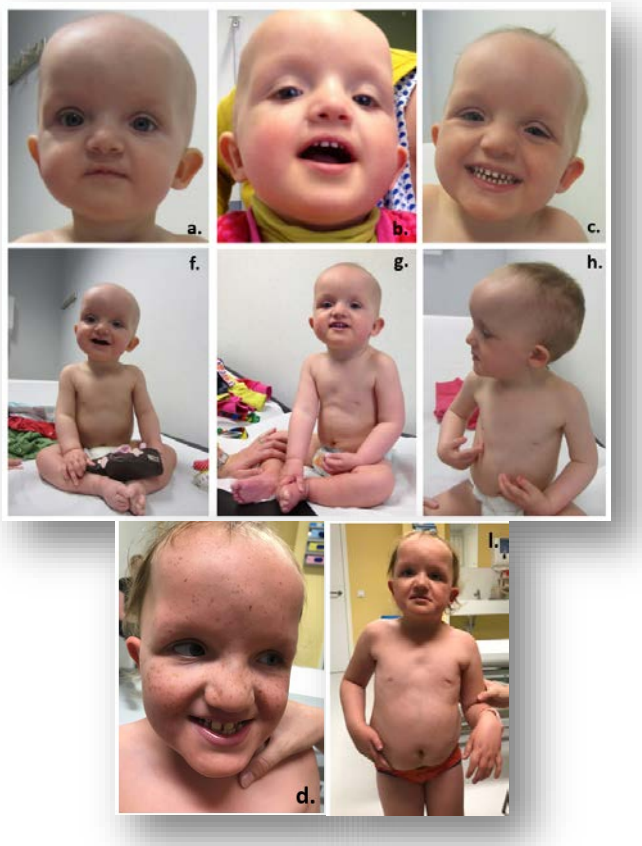
MCA/ID: trio-based WGS analysis of 3 patients with unknown clinical diagnosis

Patient 2

- Hypotonia
- Failure to thrive
- Progressive macrocephaly (H&W at p3, OFC at p97)
- Periventricular leukomalacia on imaging
- Epilepsy
- Frontal b
- Deep-se
- Downsla
- Mild hyp
- Mild intel
- Clear pic
  - Ve
  - Mil
  - Small and fragile teeth

Molecular analysis:

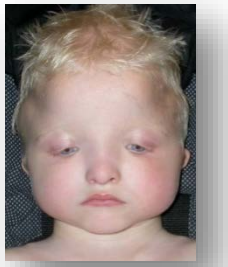
- Array CGH
- Fragile X
- PTEN, MID1 and NEMO genes
- gene panel for Rasopathies (PTPN11, SOS1, RAF1, RIT1, KRAS, BRAF, MEK1, MEK2 and HRAS)
- Mendeliome in 2015



MCA/ID: trio-based WGS analysis of 3 patients with unknown clinical diagnosis

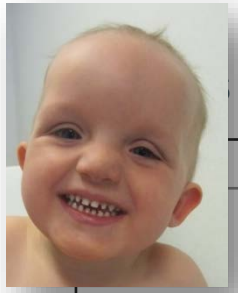
SNVs Analysis:

	Number of variants	
Filtering criteria	**Patient 1**	**Patient 2**
Total	108071	105933
Not in negative panel	104375	102349
Not in-house class 1-2	102901	100890
Frequency < 2%	8337	8083
Exonic and splice site variants	1450	1351
x-linked + recessive + de novo	95	75
AD filtering	60	58



Negative

MCA/ID: trio-based WGS analysis of 3 patients with unknown clinical diagnosis



Analysis:

Filtering criteria	Number of variants Patient 1	Patient 2
Total	108071	105933
Not in negative panel	104375	102349
Not in-house class 1-2	102901	100890
Frequency < 1%	8337	8083
Exonic and splice site variants	1450	1351
x-linked + recessive + de novo	95	75

de novo

c.430dupG - p.Ala144Glyfs*52 in ***HIST1H1E***

HGMD accession	Reported disease/phenotype	Variant class	Gene symbol	I
CI176502	Intellectual disability with overgrowth	DM	HIST1H1E	GCGACG^GG

Literature citation	Citation type	Support	
1. Tatton-Brown (2017) *Am J Hum Genet* **100:** 725 PubMed: 28475857 Mutations in Epigenetic Regulation Genes Are a Major Cause of Overgrowth with Intellectual Disability.	Primary literature report	☑	See Table S1.

Extra information	
Coding strand genomic sequence (GRCh38)	CGGCGAAGAAGCCCAAGAAGGCGACGGGGG(-/g)CGGCC
Genomic coordinate (GRCh38)	chr6:26156820-26156821
Genome viewers	UCSC; NCBI MapViewer; NCBI SeqViewer
HGVS nomenclature	NM_005321.2: c.430dupG; NP_005312.1: p.(Ala144Glyfs*52)

HISTONE GENE CLUSTER 1, H1 HISTONE FAMILY, MEMBER E; HIST1H1E

OMIM®

HGNC Approved Gene Symbol: H1-4
Cytogenetic location: 6p22.2
Genomic coordinates (GRCh38): 6:26,156,330-26,157,114 (from NCBI)

Location	Phenotype	Phenotype MIM number	Inheritance
6p22.2	**Rahman syndrome**	617537	AD

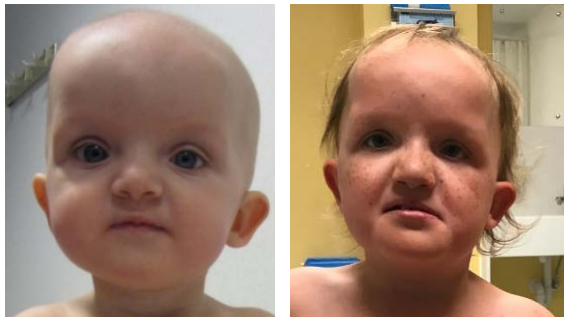
MCA/ID: trio-based WGS analysis of 3 patients with unknown clinical diagnosis

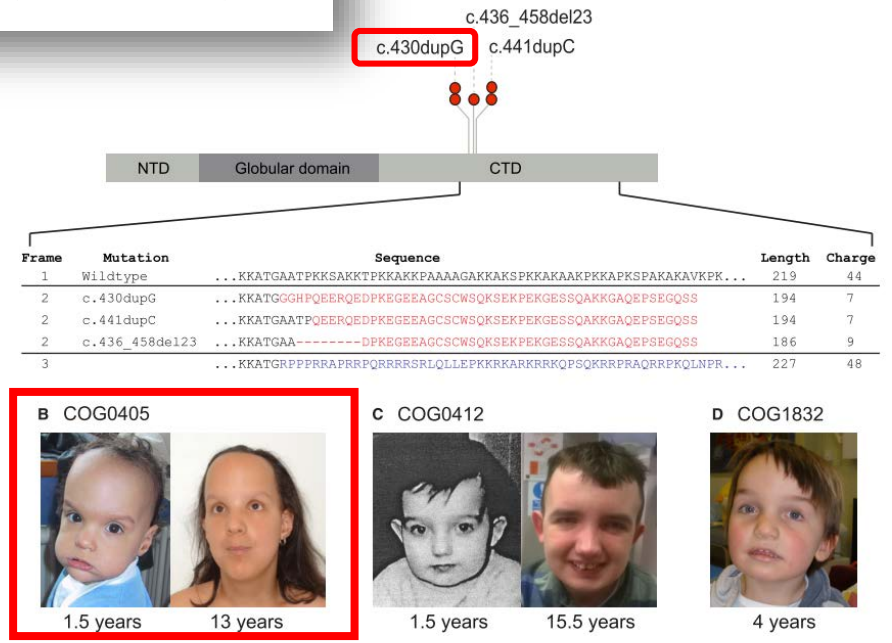
ARTICLE

Mutations in Epigenetic Regulation Genes Are a Major Cause of Overgrowth with Intellectual Disability

Katrina Tatton-Brown,[1,2] Chey Loveday,[1] Shawn Yost,[1] Matthew Clarke,[1] Emma Ramsay,[1] Anna Zachariou,[1] Anna Elliott,[1] Harriet Wylie,[1] Anna Ardissone,[3] Olaf Rittinger,[4] Fiona Stewart,[5] I. Karen Temple,[6,7] Trevor Cole,[8] Childhood Overgrowth Collaboration, Shazia Mahamdallie,[1] Sheila Seal,[1] Elise Ruark,[1] and Nazneen Rahman[1,9,10,*]

c.430dupG p.Ala144Glyfs*52



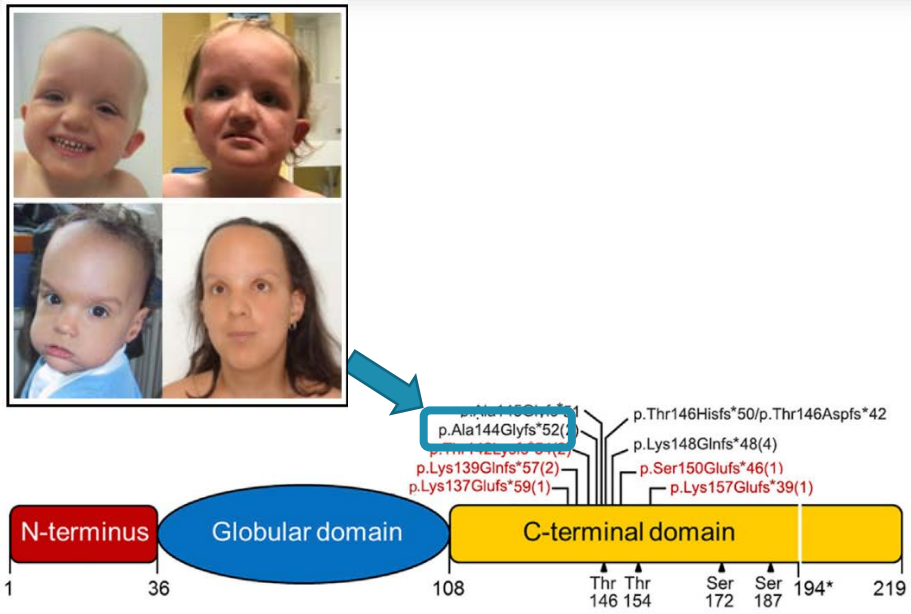


MCA/ID: trio-based WGS analysis of 3 patients with unknown clinical diagnosis

Am J Hum Genet. 2019 Sep 5;105(3):493-508. doi: 10.1016/j.ajhg.2019.07.007. Epub 2019 Aug 22.

Aberrant Function of the C-Terminal Tail of HIST1H1E Accelerates Cellular Senescence and Causes Premature Aging.

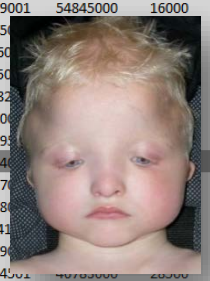
Flex E[1], Martinelli S[2], Van Dijck A[3], Ciolfi A[4], Cecchetti S[5], Coluzzi E[6], Pannone L[7], Andreoli C[8], Radio FC[4], Pizzi S[4], Carpentieri G[7], Bruselles A[2], Catanzaro G[9], Pedace L[10], Miele E[10], Carcarino E[11], Ge X[12], Chijiwa C[13], Lewis MES[13], Meuwissen M[14], Kenis S[15], Van der Aa N[14], Larson A[16], Brown K[16], Wasserstein MP[17], Skotko BG[18], Begtrup A[19], Person R[19], Karayiorgou M[20], Roos JL[21], Van Gassen KL[22], Koopmans M[22], Bijlsma EK[23], Santen GWE[23], Barge-Schaapveld DQCM[23], Ruivenkamp CAL[23], Hoffer MJV[23], Lalani SR[24], Streff H[24], Craigen WJ[24], Graham BH[25], van den Elzen APM[26], Kamphuis DJ[27], Õunap K[28], Reinson K[28], Pajusalu S[29], Wojcik MH[30], Viberti C[31], Di Gaetano C[31], Bertini E[4], Petrucci S[32], De Luca A[33], Rota R[10], Ferretti E[34], Matullo G[31], Dallapiccola B[4], Sgura A[6], Walkiewicz M[35], Kooy RF[36], Tartaglia M[37].



MCA/ID: trio-based WGS analysis of 3 patients with unknown clinical diagnosis

CNVs Analysis: → **Negative**

SVs Analysi



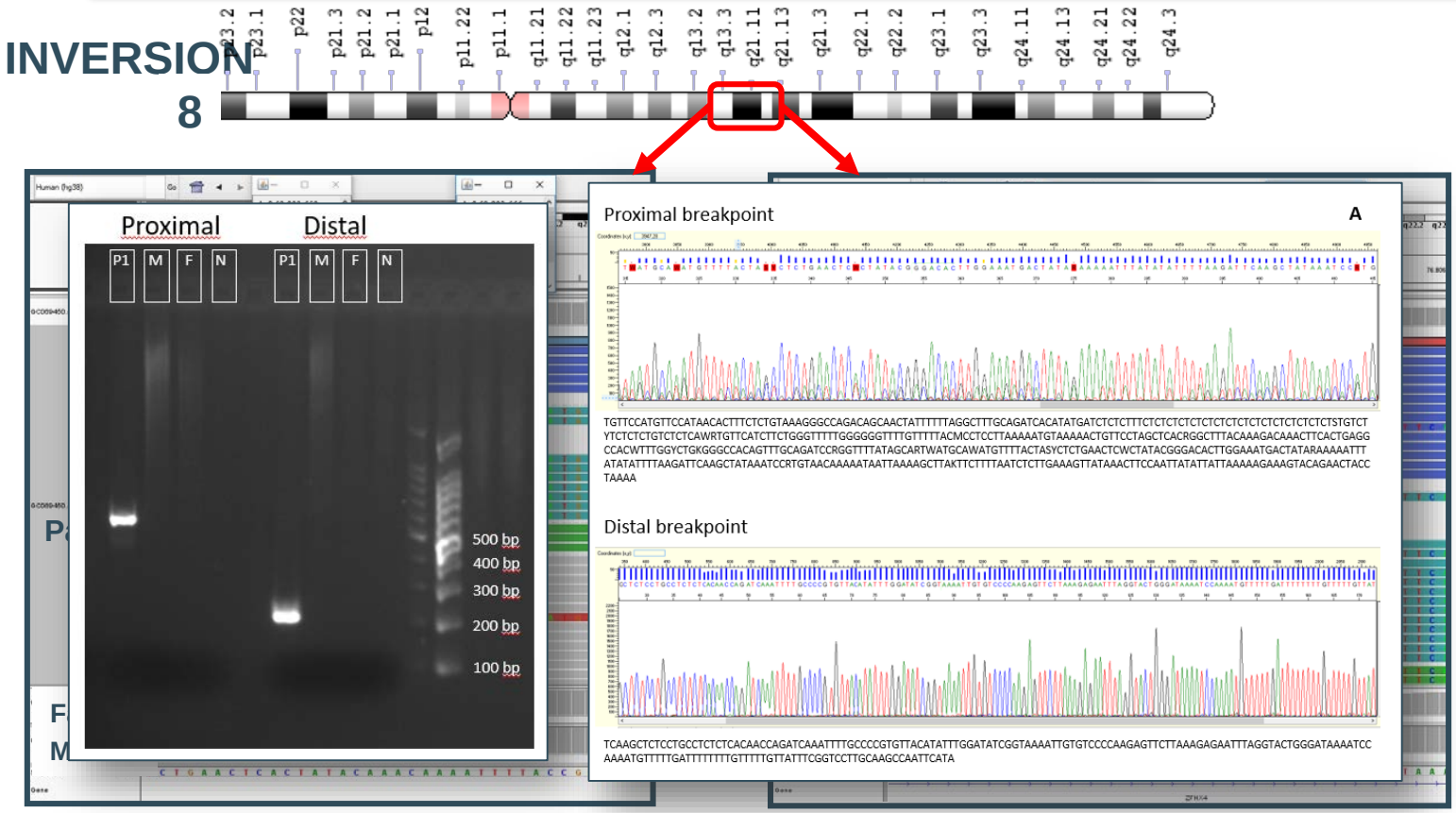
	Number of variants		
Total	20414		
Filtering criteria	**DEL/DUP/INV**	**INS**	**BND**
Good quality and rare	737	3795	2827
> 10 bp	*	165	*
De novo and recessive containing gene and/or regulatory regions	44	20	40
IGV manual filtering	3	4	0

Inversion

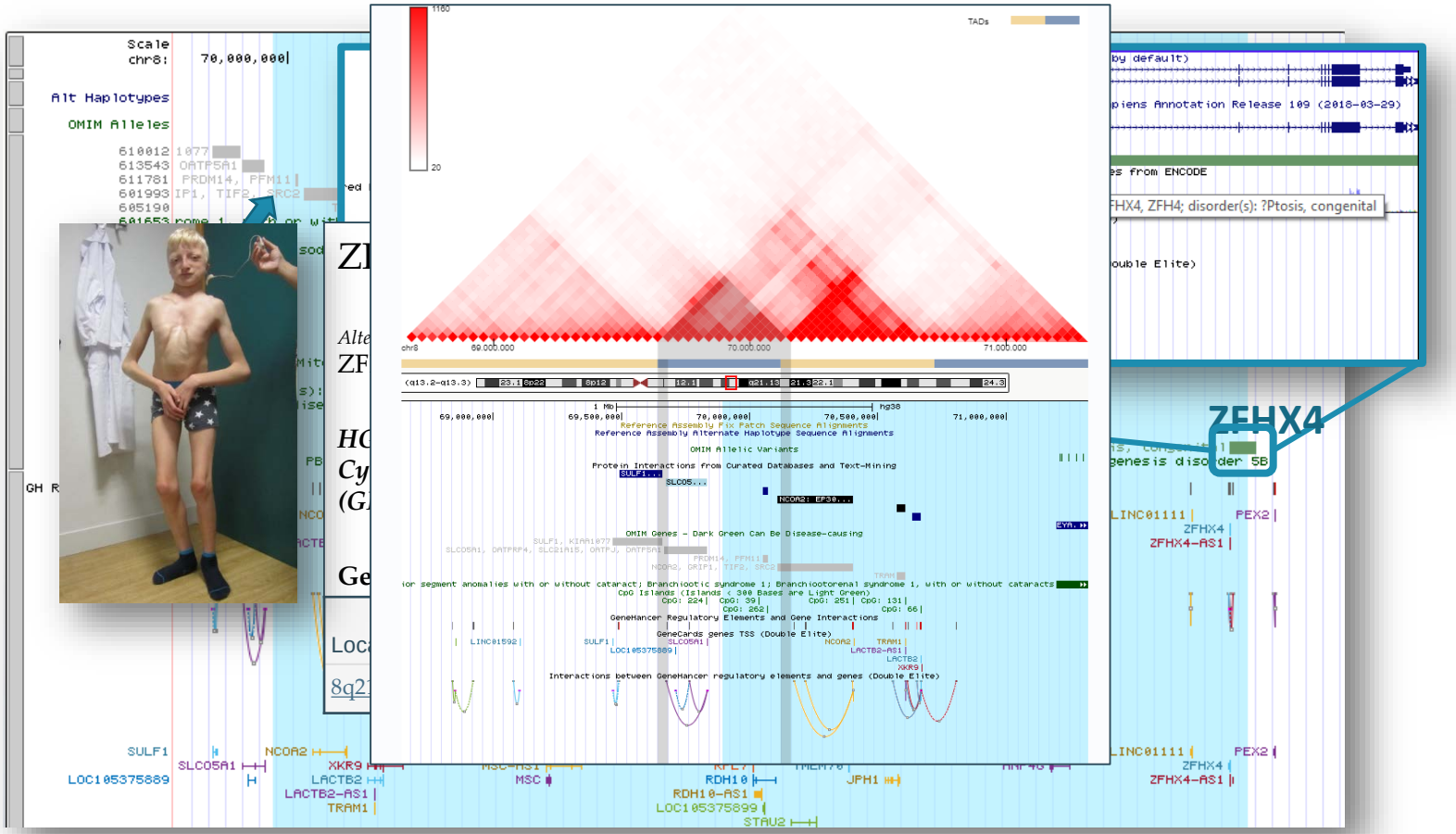
chr8:69893659-76806725

6,9 Mb

MCA/ID: trio-based WGS analysis of 3 patients with unknown clinical diagnosis



MCA/ID: trio-based WGS analysis of 3 patients with unknown clinical diagnosis



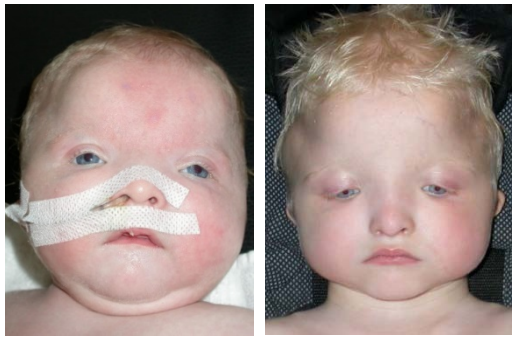
MCA/ID: trio-based WGS analysis of 3 patients with unknown clinical diagnosis

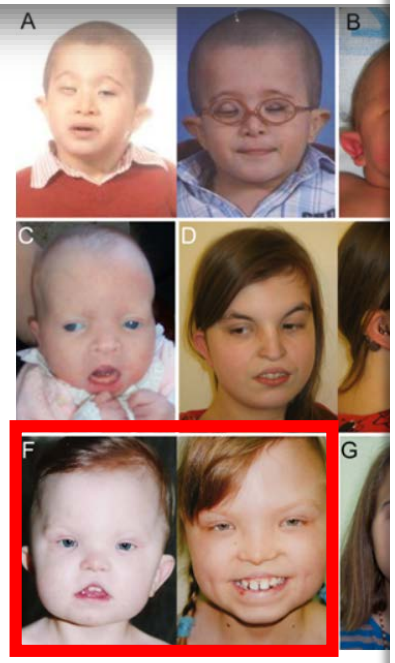
Am J Hum Genet. 2011 Aug 12;89(2):295-301. doi: 10.1016/j.ajhg.2011.06.012. Epub 2011 Jul 28.

Characterization of a 8q21.11 microdeletion syndrome associated with i recognizable phenotype.

Palomares M[1], Delicado A, Mansilla E, de Torres ML, Vallespín E, Fernandez L, Martinez-Glez V, García-Miñaur S, N Lynch SA, Sharkey FH, Thuresson AC, Annerén G, Belligni EF, Martínez-Fernández ML, Bermejo E, Nowakowska B Obersztyn E, Martínez-Frías ML, Hennekam RC, Lapunzina P.

GENEMATCHER





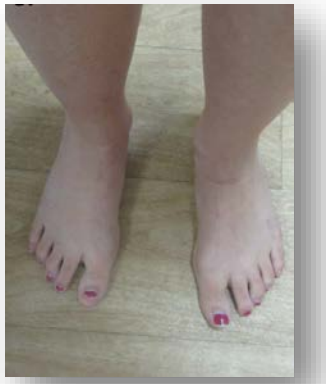
- Round face with full cheeks
- High forehead
- Ptosis
- Corneal opacities
- Wide nasal bridge
- Underdeveloped alae
- Short philtrum
- Cupid's bow of the upper lip
- Downturned corners of the mouth
- Micrognathia
- Low-set and prominent ears
- Short neck

- Camptodactyly
- Syndactyly
- Broadening of the first rays
- Hypotonia
- Impaired balance
- Sensorineural hearing loss
- Underdeveloped corpus callosum
- Unusual behavior

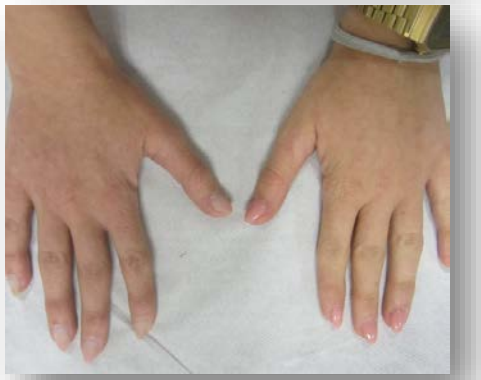
MCA/ID: trio-based WGS analysis of 3 patients with unknown clinical diagnosis

22 years old

- Non-consanguineous, healthy parents
- Ventricular septum defect
- Coarctation of the aorta
- Horseshoe kidney
- Bilateral choanal atresia
- Clinodactyly of the third and fourth finger
- Bilateral sandal gap
- Short stature
- Hyperextension of the knees and slumped shoulders
- Hypogenesis of the abdominal mesentery
- Mild intellectual disability
- Facial dysmorphism
  - Midfacial hypoplasia
  - Short palpebral fissures
  - High-arched palate
  - Undersized maxilla resulting in a nasal speech
  - Ptosis of the upper eyelids
  - Smallmouth and ears
  - Horner's syndrome



Molecular analysis:
- CHD7 negative
- Array-CGH negative



MCA/ID: trio-based WGS analysis of 3 patients with unknown clinical diagnosis

SNVs Analysis:

Filtering criteria	*Number of variants*
Total	6,912,472
De novo variants	102,190
Rare variants (MAF<1%)	69,071
Exonic and splice-site variants	223
CADD > 20	142
Excluding synonymous variants	91

→ **Negative**

MCA/ID: trio-based WGS analysis of 3 patients with unknown clinical diagnosis

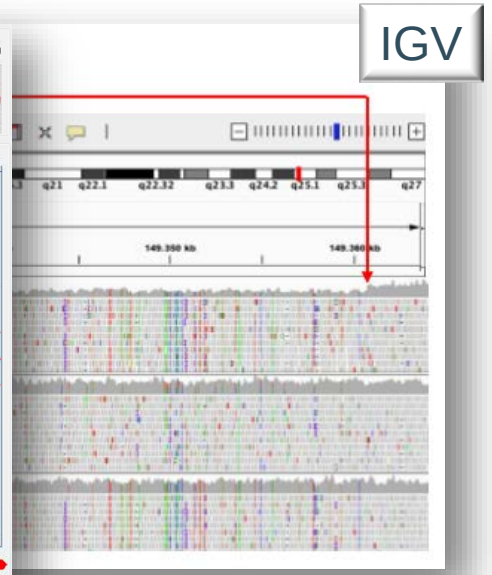
CNVs Analysis:

- fold change under 0.7 and above 1.3
- good mappability
- *de novo*

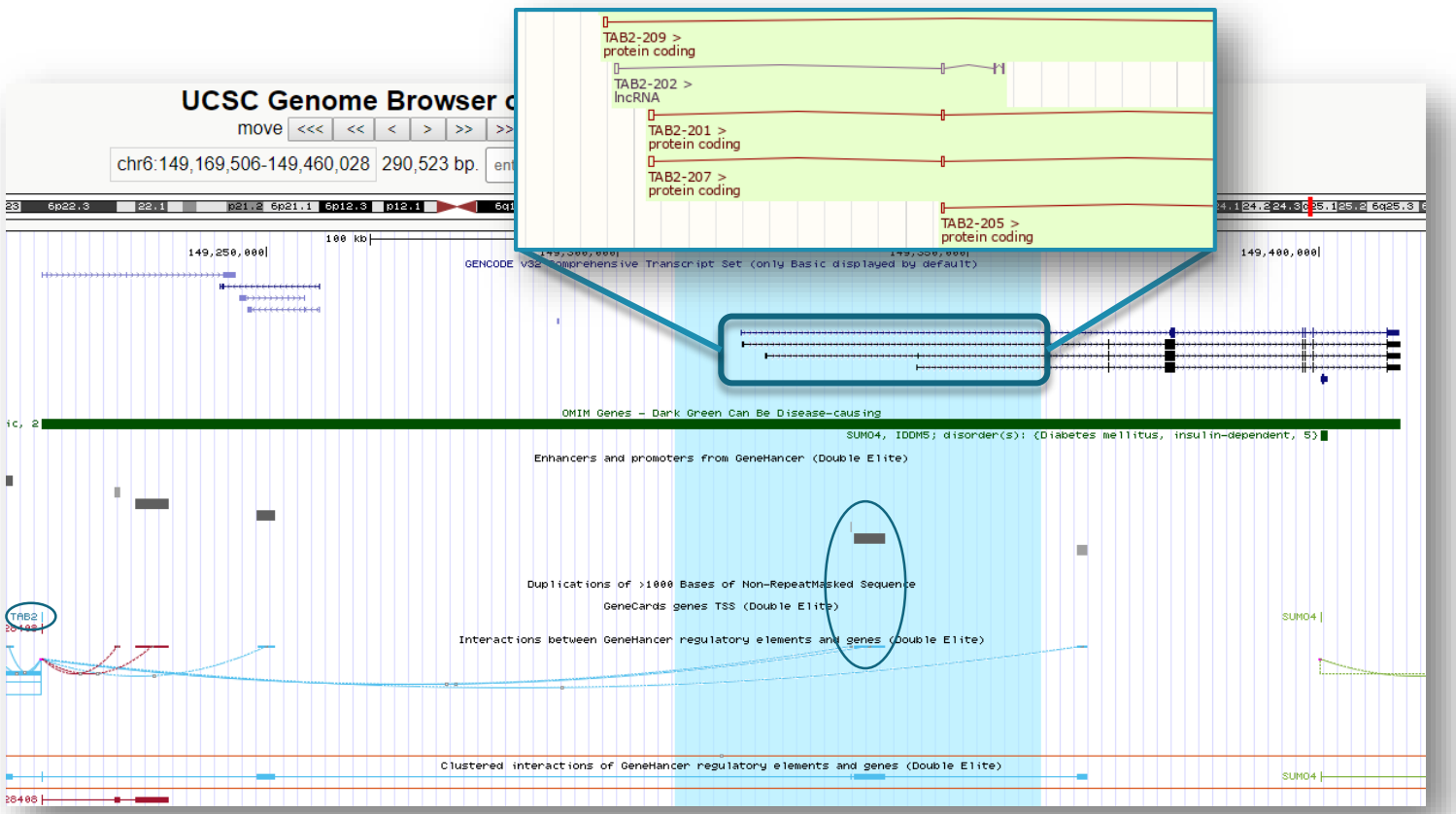
deletion
chr6: 149,308,196 - 149,360,335
in *TAB2* gene

Array-CGH

IGV



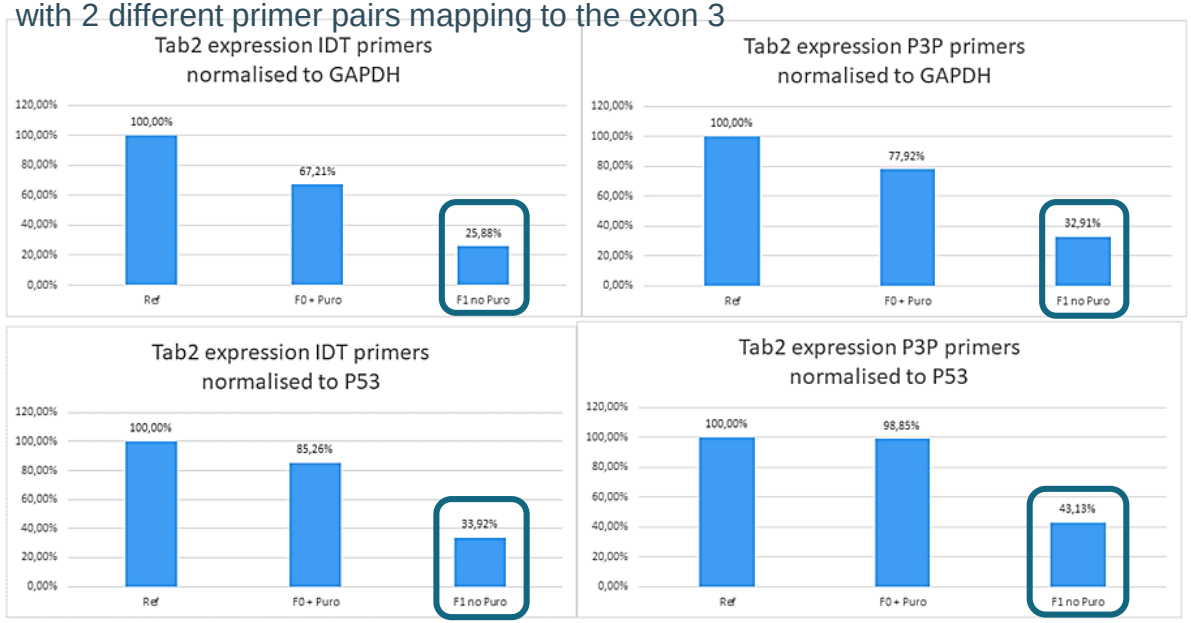
MCA/ID: trio-based WGS analysis of 3 patients with unknown clinical diagnosis



Gene expression

qRT-PCR

with 2 different primer pairs mapping to the exon 3



MCA/ID: trio-based WGS analysis of 3 patients with unknown clinical diagnosis

Whole Exome Sequencing, Familial Genomic Triangulation, and Systems Biology Converge to Identify a Novel Nonsense Mutation in *TAB2*-encoded TGF-beta Activated Kinase 1 in a Child with Polyvalvular Syndrome

Jaeger P. Ackerman, BA,* John A. Smestad, BS,† David J. Tester, BS,* Muhammad Y. Qureshi, MBBS,* Beau A. Crabb, MS, CGC,‡ Nancy J. Mendelsohn, MD,‡ and Michael J. Ackerman, MD, PhD*

A recognizable systemic connective tissue disorder with polyvalvular heart dystrophy and dysmorphism associated with *TAB2* mutations

M. Ritelli[1] | S. Morlino[2] | E. Giacopuzzi[1] | L. Bernardini[3] | B. Torres[3] | G. Santoro[1] | V. Ravasio[1] | N. Chiarelli[1] | D. D'Angelantonio[2] | A. Novelli[4] | P. Grammatico[2] | M. Colombi[1] | M. Castori[5]

A 2.6 Mb deletion of 6q24.3–25.1 in a patient with growth failure, cardiac septal defect, thin upper lip and asymmetric dysmorphic ears

R. Caselli [a], M.A. Mencarelli [a], F.T. Papa [a], V. Uliana [a], S. Schiavone [b], M. Strambi [b], C. Pescucci [a], F. Ariani [a], V. Rossi [c], I. Longo [a], I. Meloni [a], A. Renieri [a,*], F. Mari [a]

Familial *TAB2* Microdeletion and Congenital Heart Defects Including Unusual Valve Dysplasia and Tetralogy of Fallot

Karin Weiss,[1] Carolyn Applegate,[2] Tao Wang,[2,3] and Denise A. S. Batista[2,4,5]*

A De Novo 0.63 Mb 6q25.1 Deletion Associated with Growth Failure, Congenital Heart Defect, Underdeveloped Cerebellar Vermis, Abnormal Cutaneous Elasticity and Joint Laxity

Vincenzo Salpietro,[1,2]* Martino Ruggieri,[3] Kshitij Mankad,[4] Gabriella Di Rosa,[5] Francesca Granata,[6] Italia Loddo,[2] Emanuela Moschella,[2] Maria Pia Calabro,[7] Anna Capalbo,[8] Laura Bernardini,[8] Antonio Novelli,[9] Agata Polizzi,[10,11] Daniela G. Seidler,[12] Teresa Arrigo,[2] and Silvana Briuglia[2]

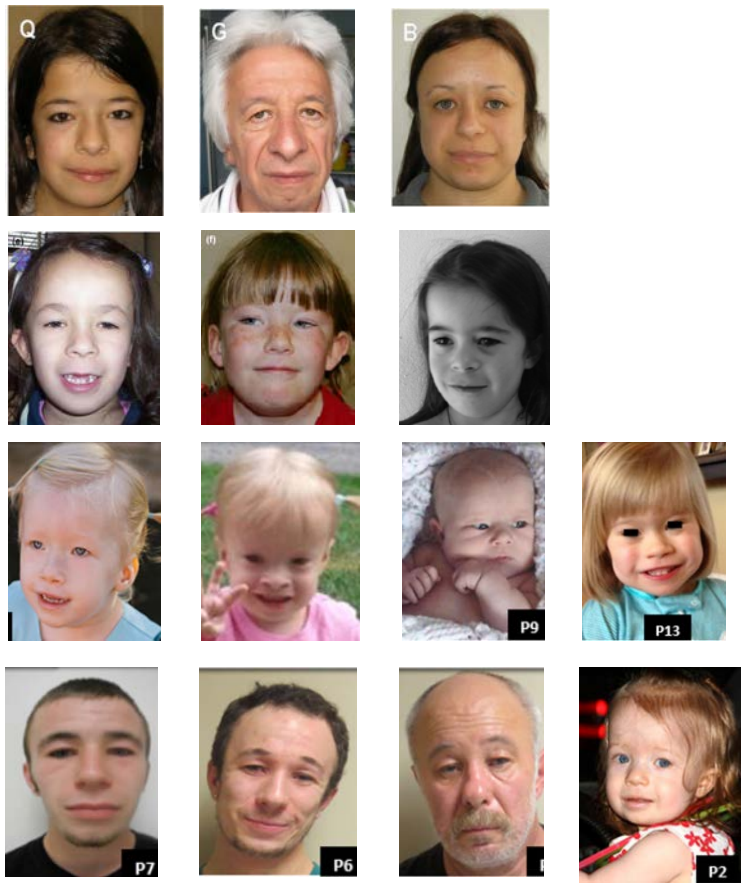
Haploinsufficiency of *TAB2* Causes Congenital Heart Defects in Humans

Bernard Thienpont,[1,14] Litu Zhang,[2,15] Alex V. Postma,[3] Jeroen Breckpot,[1] Léon-Charles Tranchevent,[4] Peter Van Loo,[5,6] Kjeld Møllgård,[7] Niels Tommerup,[2] Iben Bache,[2] Zeynep Tümer,[2,8] Klaartje van Engelen,[9] Björn Menten,[10] Geert Mortier,[10,11] Darrel Waggoner,[12] Marc Gewillig,[13] Yves Moreau,[4] Koen Devriendt,[1] and Lars Allan Larsen[2,*]

6q25.1 (*TAB2*) microdeletion syndrome: Congenital heart defects and cardiomyopathy

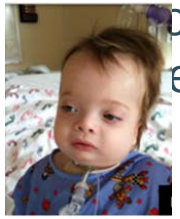
Andrew Cheng[1] | Mary Beth P. Dinulos[2] | Whitney Neufeld-Kaiser[3] | Jill Rosenfeld[4] | McKenna Kyriss[5] | Suneeta Madan-Khetarpal[6] | Hiba Risheg[7] | Peter H. Byers[3] | Yajuan J. Liu[3]

MCA/ID: trio-based WGS analysis of 3 patients with unknown clinical diagnosis

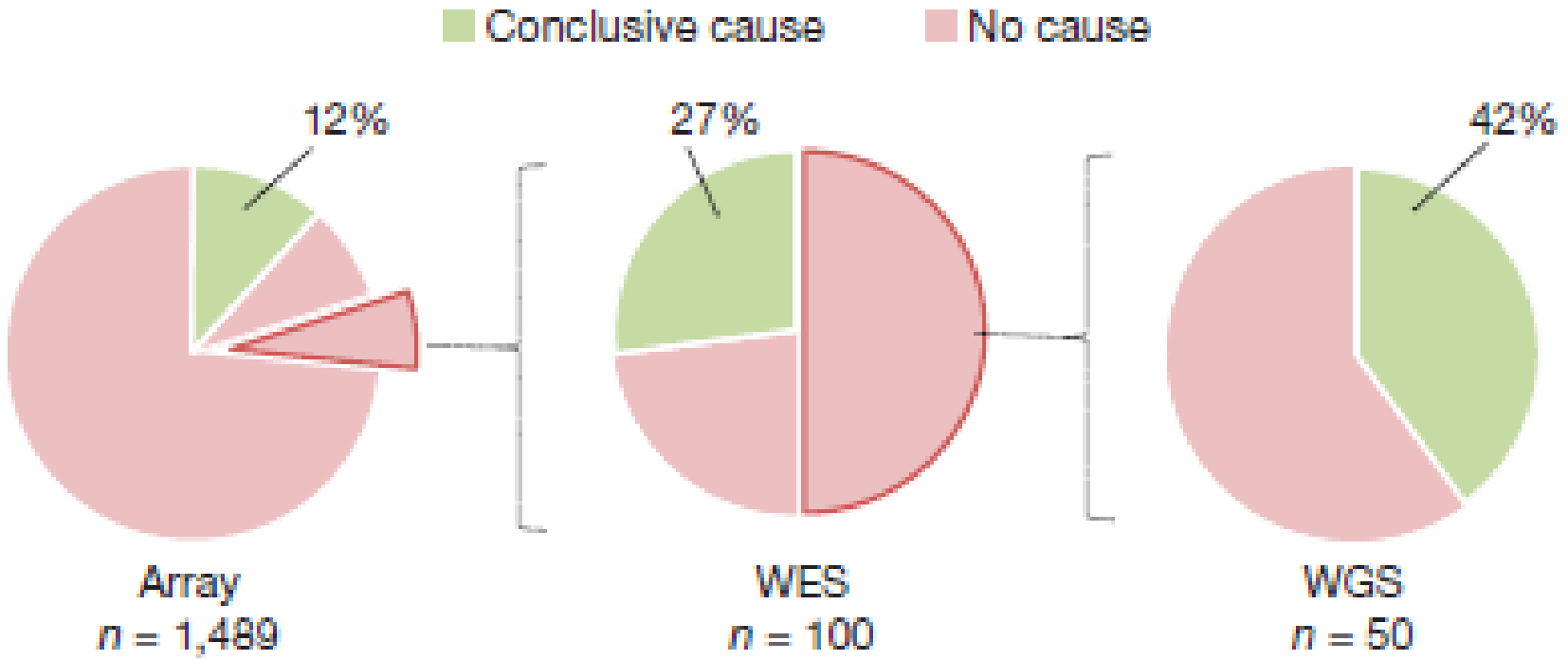


- CHDs
- Facial dysmorphism
- Growth failure
- Joint laxity
- Hypotonia
- Connective tissue abnormalities
- Developmental or intellectual disability

- Horseshoe kidney
- Bilateral choanal atres
- ogenesis of the ab
- entery

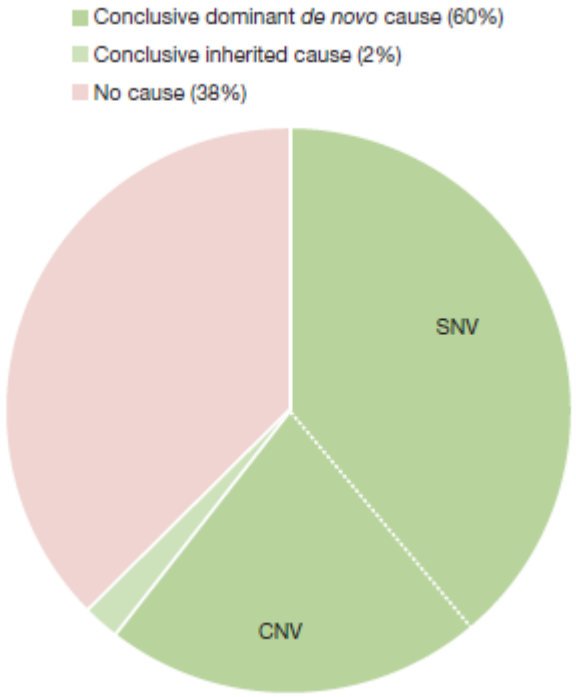


Analysis of the complete genome (SNPs + CNVs)



Genome sequencing identifies major causes of severe intellectual disability

Analysis of the complete genome



Diagnostic yield: 62%

Genome sequencing identifies major causes of severe intellectual disability

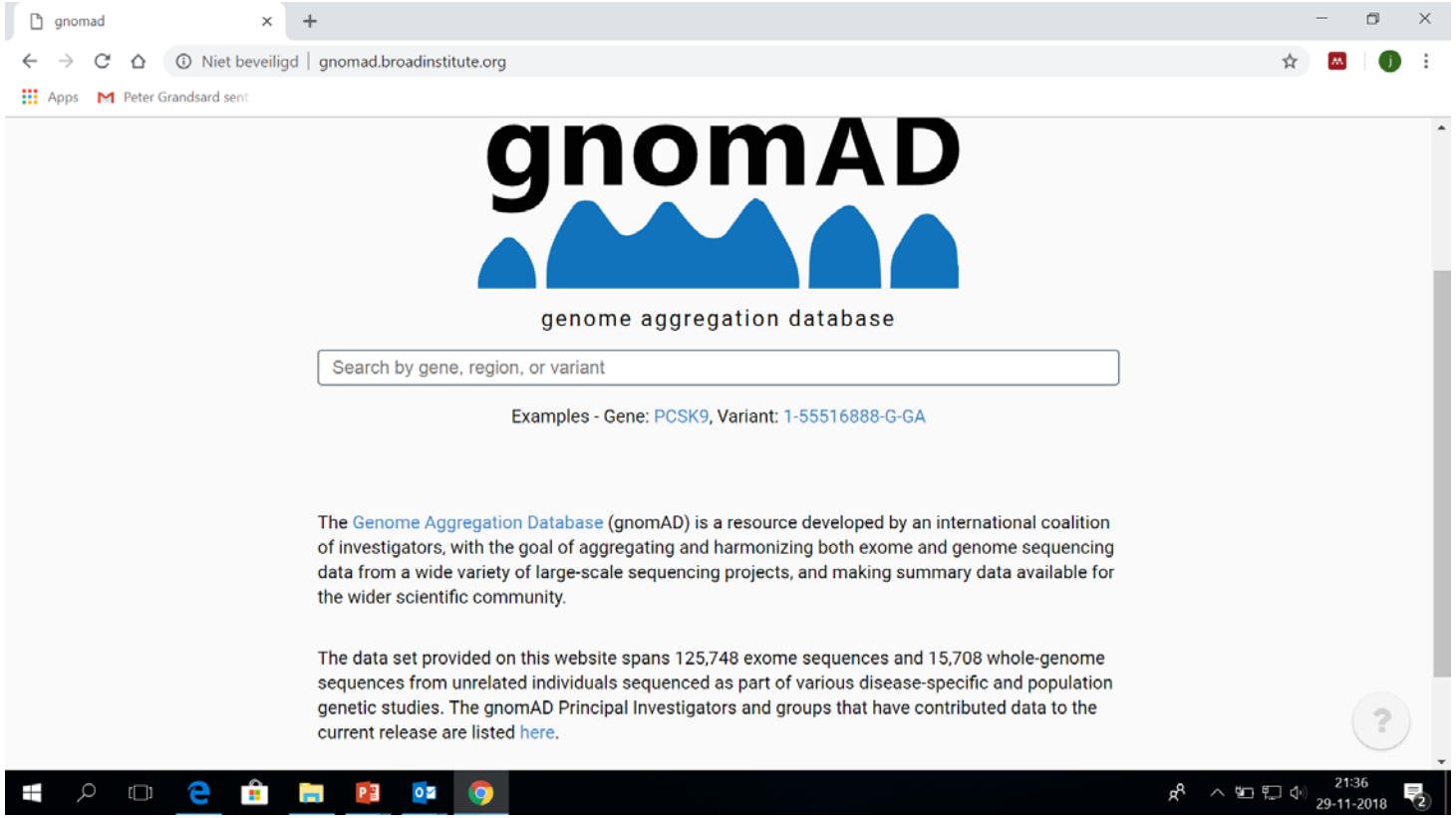
Christian Gilissen[1*], Jayne Y. Hehir-Kwa[1*], Djie Tjwan Thung[1], Maartje van de Vorst[1], Bregje W. M. van Bon[1], Marjolein H. Willemsen[1], Michael Kwint[1], Irene M. Janssen[1], Alexander Hoischen[1], Annette Schenck[1], Richard Leach[2], Robert Klein[2], Rick Tearle[2], Tan Bo[1,3], Rolph Pfundt[1], Helger G. Yntema[1], Bert B. A. de Vries[1], Tjitske Kleefstra[1], Han G. Brunner[1,4*], Lisenka E. L. M. Vissers[1*] & Joris A. Veltman[1,4*]

The burden of autosomal recessive diseases in rare developmental disorders

- DDD study
  - 3.6% autosomal recessive
  - 40% de novo coding mutations
- Pakistani study:
  - 30.9% autosomal recessive
  - 30% de novo dominant

gnomAD (past Exac) database

exercise



We are all mutants!