Pathogenic alleles, clan genomics and the complex architecture of human disease

"High-throughput Sequencing – Applications and Analyses
Oslo University 200th, NORWAY
October 28, 2011

James R. Lupski, M.D., Ph.D., D.Sc.
Department of Molecular & Human Genetics
& Department of Pediatrics
Baylor College of Medicine
& Texas Children's Hospital
Houston, TX

Medical Genetics Laboratories
Department of Molecular and Human Genetics • Houston, Texas
http://www.bcm.edu/geneticlabs/

Disclosure

Texas Children's Hospital

J.R.L. is a consultant for:

Athena Diagnostics

23andMe

Ion Torrent Systems Inc.

NO affiliation with:
Life Technologies, Inc
NOR
Applied BioSysytems, Inc

Co-Inventor on Diagnostic Patents:

UNITED STATES: 5,294,533 (issued 03/15/94); 5,306,616 (issued 04/26/94); 5,523,217 (issued 06/04/96); 5,599,920 (issued 02/04/97); 5,667,968 (issued 09/16/97);5,780,223 (issued 07/14/98); 6,132,954 (issued 10/19/00); 6,713,300 (issued 03/30/04); 7,141,420 (issued 11/28/06); 7,189,511 (issued 03/13/07); 7,192,579 (issued 03/20/07); 7,273,698 (issued 09/25/07).

EUROPEAN: 0424473 (issued 05/08/96), 0610396 (issued 01/17/01), 0989805 (issued 01/11/06).

The Medical Genetics Laboratories (MGL) of the Dept of Molecular and Human Genetics at Baylor College of Medicine derives revenue from molecular diagnostic testing. MGL, http://www.bcm.edu/geneticlabs/

BCM Baylor College of Medicine — Medical Genetics Laboratories — Department of Molecular and Human Genetics • Houston, Texas

BCM Baylor College of Medicine

A story about Charcot-Marie-Tooth disease

CMT: clinical & genetic aspects

The CMT1A duplication
- a paradigm for CNV mutation & mechanisms for CNV formation

CMT mutational load
- gene load? locus load? or genomic load?
- SNP + CNV

Personal genome sequencing: CMT

Genetic contributions to inherited and apparently acquired neurologic dz

CMT: clinical & genetic aspects

The CMT1A duplication
- a paradigm for CNV mutation & mechanisms for CNV formation

CMT mutational load
- gene load? locus load? or genomic load?
- SNP + CNV

Personal genome sequencing: CMT

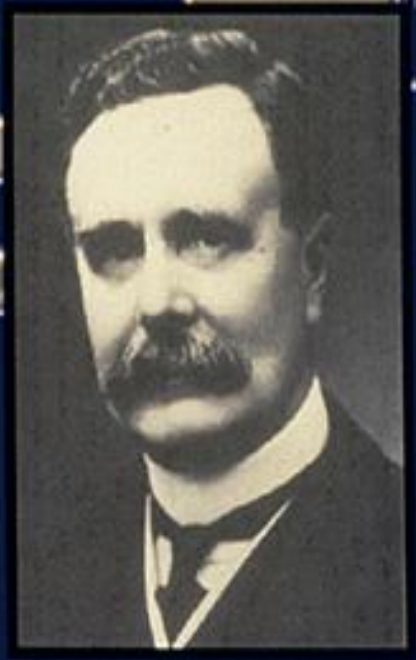
CHARCOT

1825-1893

MARIE

1853-1940

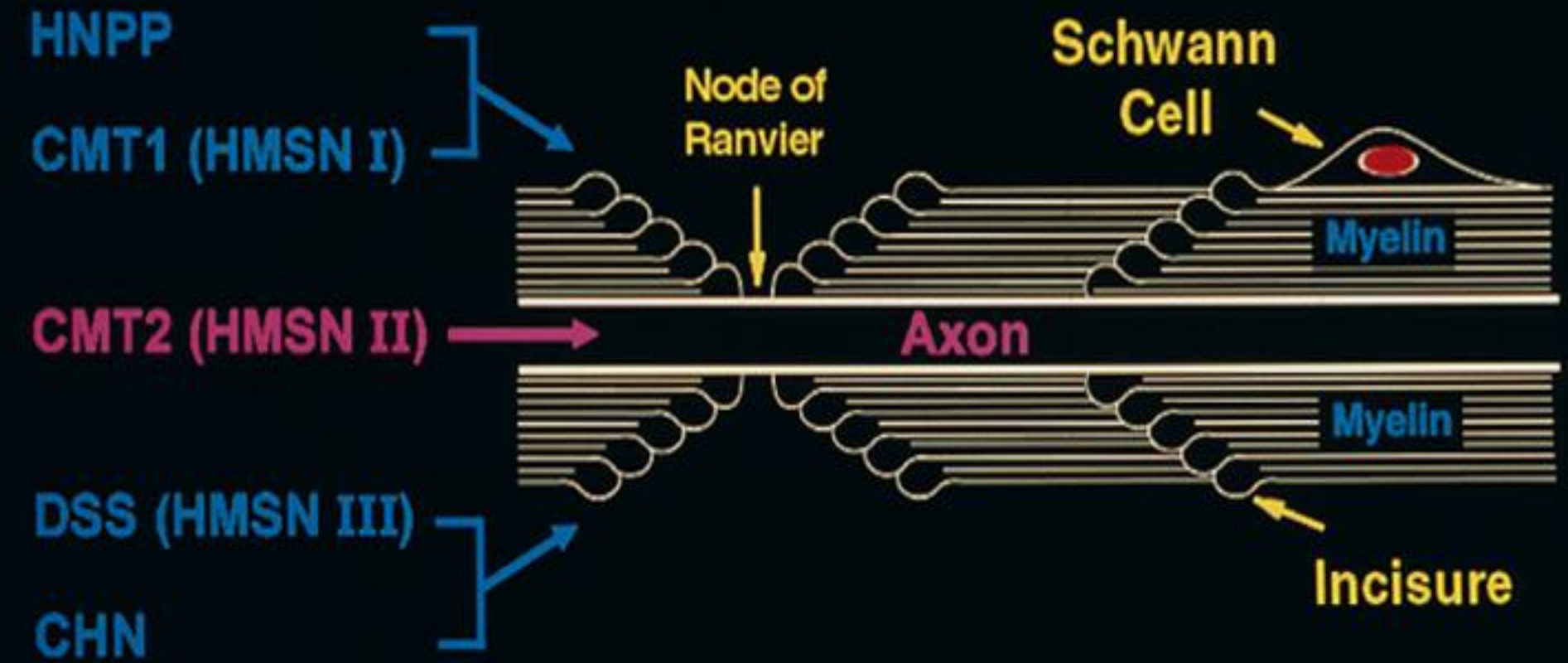
TOOTH



1856-1925

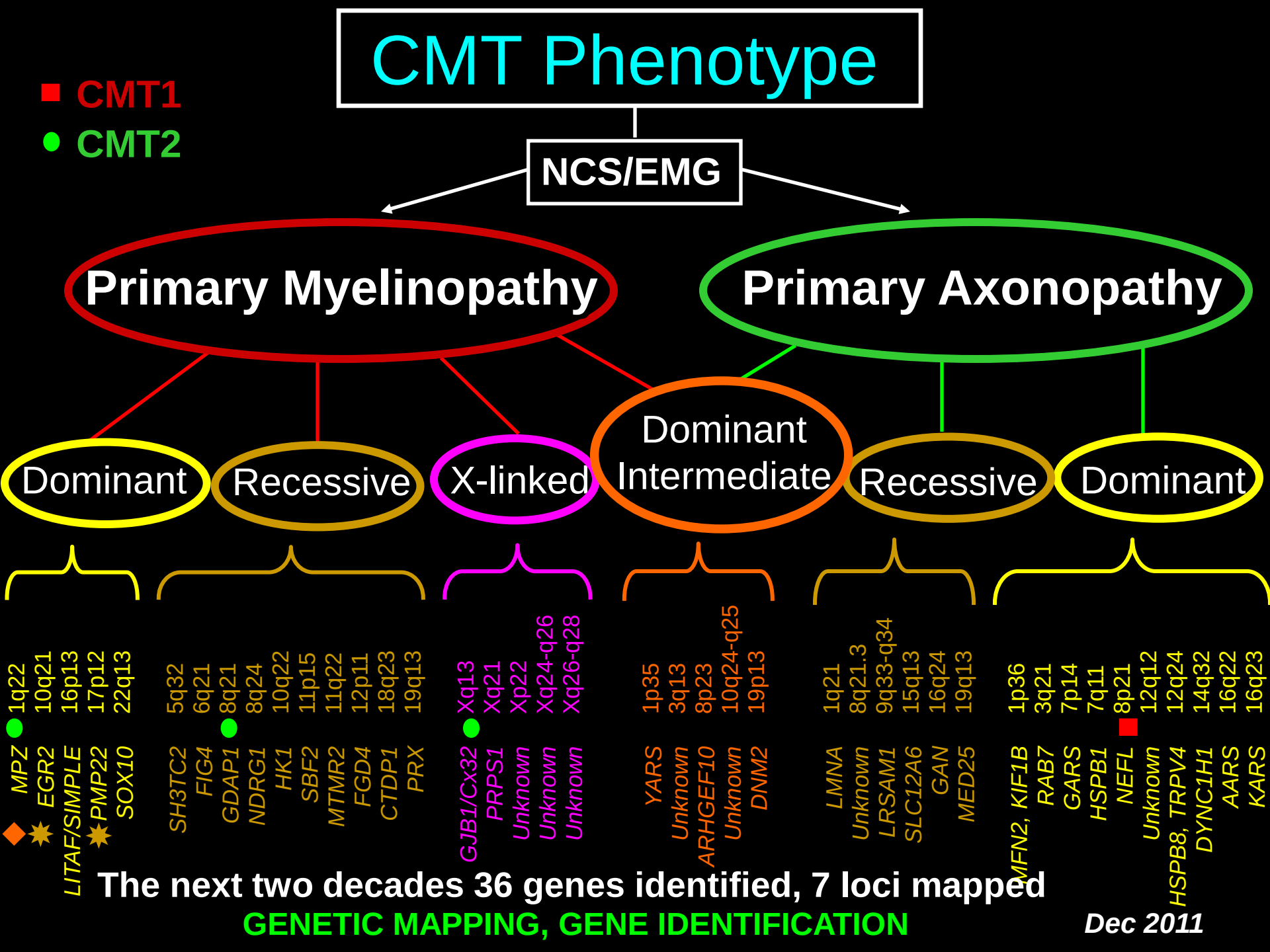
1886 in Paris & Cambridge,UK

Hereditary Neuropathies



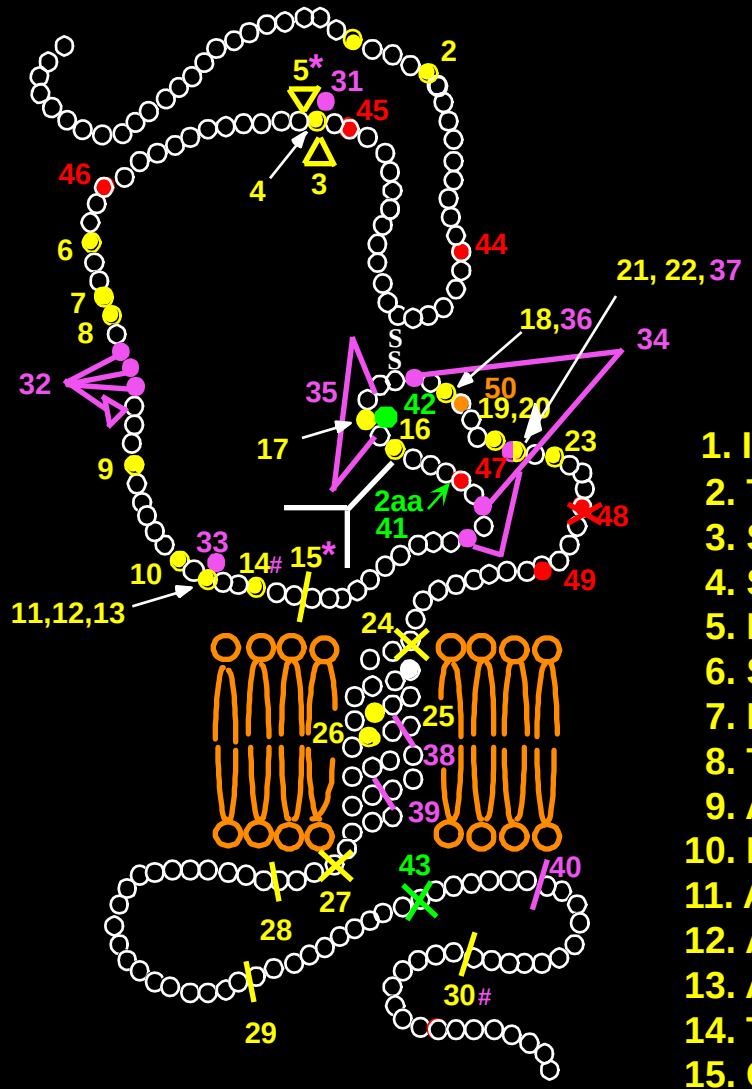
CMT research – the first century (105 years!).

CLINICAL DESCRIPTIONS


CMT Phenotype
CMT1
CMT2
NCS/EMG
Primary Myelinopathy
Primary Axonopathy
Dominant
Recessive
X-linked
Dominant Intermediate
Recessive
Dominant
1q22 MPZ
10q21 EGR2
16p13 LITAF/SIMPLE
17p12 PMP22
22q13 SOX10
5q32 SH3TC2
6q21 FIG4
8q21 GDAP1
8q24 NDRG1
10q22 HK1
11p15 SBF2
11q22 MTMR2
12p11 FGD4
18q23 CTDP1
19q13 PRX
Xq13 GJB1/Cx32
Xq21 PRPS1
Xp22 Unknown
Xq24-q26 Unknown
Xq26-q28 Unknown
1p35 YARS
3q13 Unknown
8p23 ARHGEF10
10q24-q25 Unknown
19p13 DNM2
1q21 LMNA
8q21.3 Unknown
9q33-q34 LRSAM1
15q13 SLC12A6
16q24 GAN
19q13 MED25
1p36 MFN2, KIF1B
3q21 RAB7
7p14 GARS
7q11 HSPB1
8p21 NEFL
12q12 Unknown
12q24 HSPB8, TRPV4
14q32 DYNC1H1
16q22 AARS
16q23 KARS
The next two decades 36 genes identified, 7 loci mapped
GENETIC MAPPING, GENE IDENTIFICATION
Dec 2011

Summary of P_0 Mutations

Charcot-Marie-Tooth Disease Type 1 | **Dejerine-Sottas Syndrome** | **Congenital Hypomyelination** | **Charcot-Marie-Tooth Disease Type 2** | **Roussy-Levy syndrome**

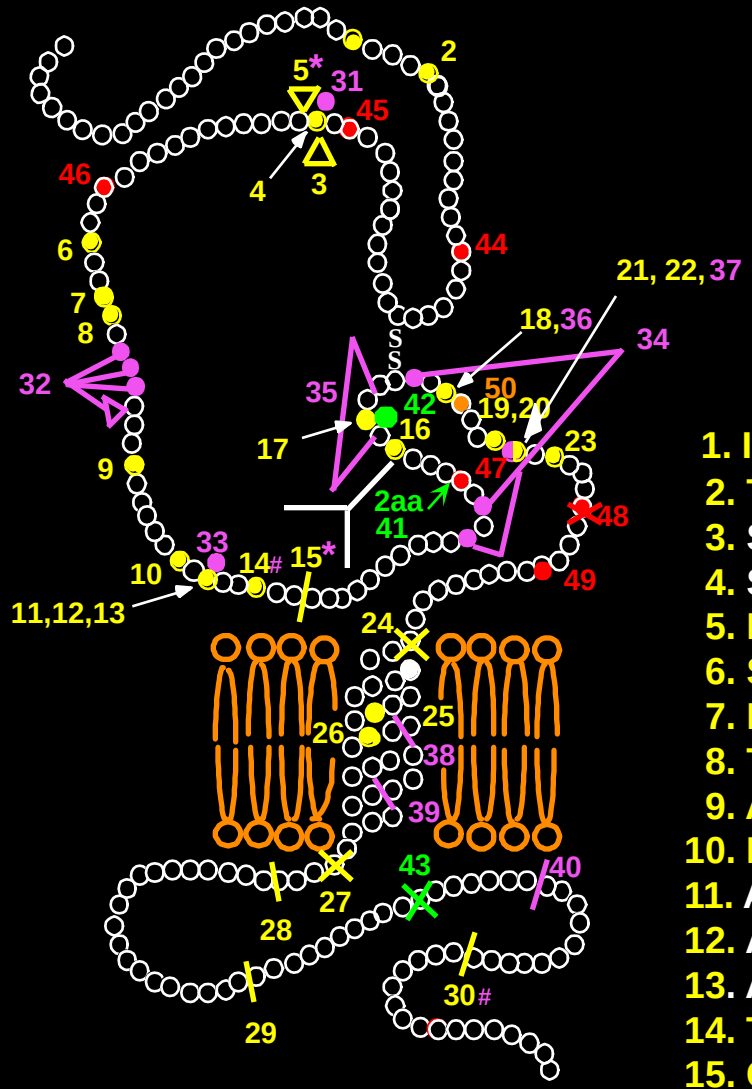


- Nonsense mutation
- Frameshift mutation
- Amino acid deletion
- Missense
- Insertion
- aa Amino acid

1. Ile(30)Met
2. Thr(34)Ile
3. Ser(63)del
4. Ser(63)Phe
5. Phe(64)del*hmz*
6. Ser(78)Leu
7. His(81)Arg
8. Tyr(82)Cys
9. Asp(90)Glu
10. Lys(96)Glu
11. Arg(98)His
12. Arg(98)Ser
13. Arg(98)Pro
14. Trp(101)Cys#
15. Gly(103)fs*hmz*
16. Asn(122)Ser
17. Thr(124)Met
18. Lys(130)Arg
19. Asp(134)Glu
20. Asp(134)Asn
21. Ile(135)Thr
22. Ile(135)Leu
23. Gly(137)Ser
24. Tyr(154)Stop
25. Gly(163)Arg
26. Gly(167)Arg
27. Tyr(181)Stop
28. Leu(184)fs
29. Lys(204)fs
30. Val(232)fs#
31. Ser(63)Cys
32. Gln(84)His, Pro(85)Leu, Tyr(86)Phe, Ile(87)del
33. Arg(98)Cys
34. Ile(114)Thr, Asn(116)His, Asp(128)Asn
35. Thr(124)del, Phe(125)del
36. Lys(130)Arg
37. Ile(135)Thr
38. Val(169)fs
39. Leu(174)fs
40. Ala(221)fs
41. Asp(118)ins2aa
42. Thr(124)Lys
43. Gln(215)Stop
44. Ser(44)Phe
45. Asp(61)Gly
46. Asp(75)Val
47. Tyr(119)Cys
48. Gln(141)Stop
49. Tyr(145)Ser
50. Asn(131)Lys

Summary of P_0 Mutations

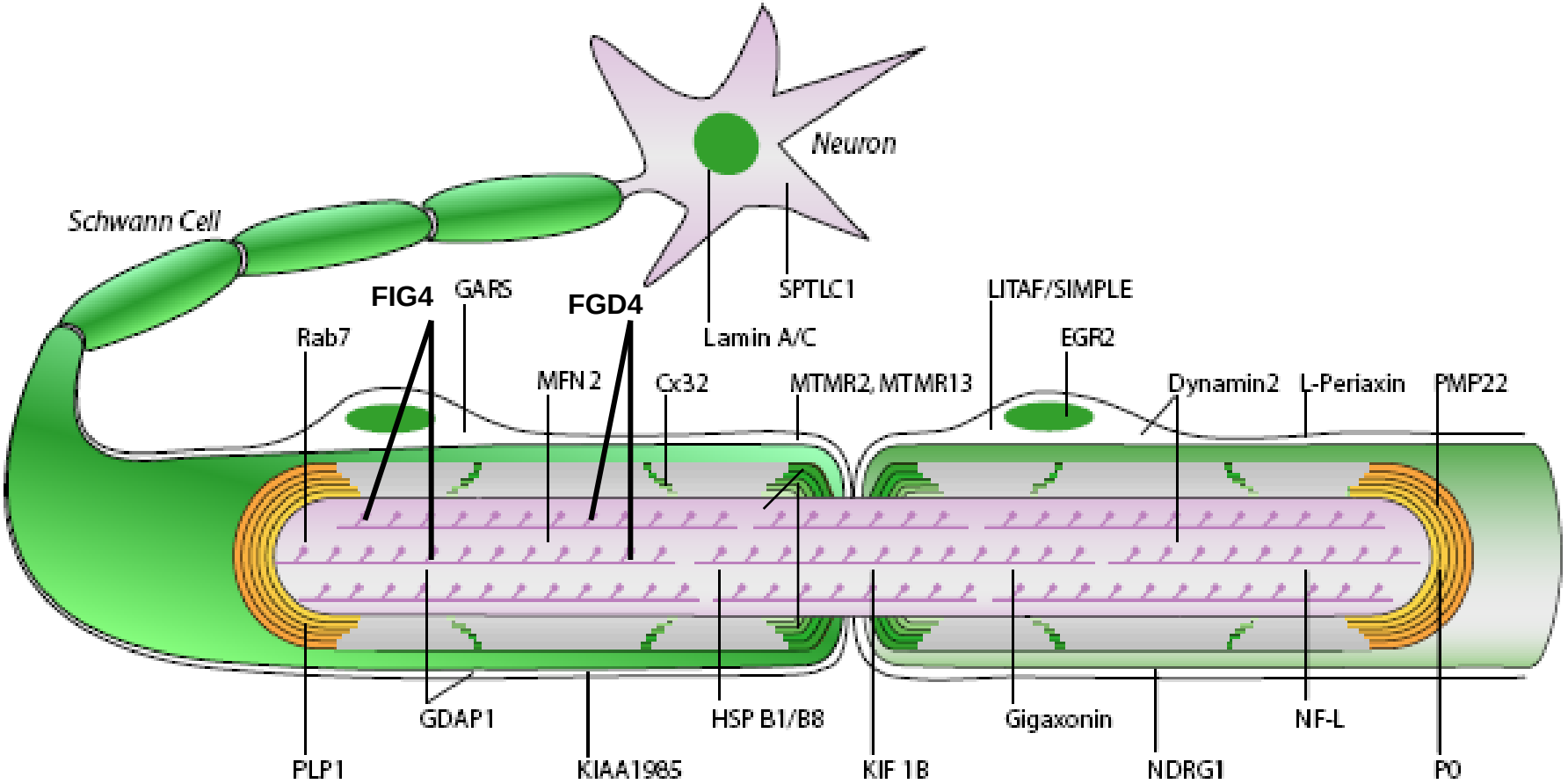
Charcot-Marie-Tooth Disease Type 1 | **Dejerine-Sottas Syndrome** | **Congenital Hypomyelination** | **Charcot-Marie-Tooth Disease Type 2** | **Roussy-Levy syndrome**



Nonsense mutation
Frameshift mutation
Amino acid deletion
Missense
Insertion
aa Amino acid

1. Ile(30)Met
2. Thr(34)Ile
3. Ser(63)del
4. Ser(63)Phe
5. Phe(64)del*hmz*
6. Ser(78)Leu
7. His(81)Arg
8. Tyr(82)Cys
9. Asp(90)Glu
10. Lys(96)Glu
11. Arg(98)His
12. Arg(98)Ser
13. Arg(98)Pro
14. Trp(101)Cys#
15. Gly(103)fs*hmz*
16. Asn(122)Ser
17. Thr(124)Met
18. Lys(130)Arg
19. Asp(134)Glu
20. Asp(134)Asn
21. Ile(135)Thr
22. Ile(135)Leu
23. Gly(137)Ser
24. Tyr(154)Stop
25. Gly(163)Arg
26. Gly(167)Arg
27. Tyr(181)Stop
28. Leu(184)fs
29. Lys(204)fs
30. Val(232)fs#
31. Ser(63)Cys
32. Gln(84)His, Pro(85)Leu, Tyr(86)Phe, Ile(87)del
33. Arg(98)Cys
34. Ile(114)Thr, Asn(116)His, Asp(128)Asn
35. Thr(124)del, Phe(125)del
36. Lys(130)Arg
37. Ile(135)Thr
38. Val(169)fs
39. Leu(174)fs
40. Ala(221)fs
41. Asp(118)ins2aa
42. Thr(124)Lys
43. Gln(215)Stop
44. Ser(44)Phe
45. Asp(61)Gly
46. Asp(75)Val
47. Tyr(119)Cys
48. Gln(141)Stop
49. Tyr(145)Ser
50. Asn(131)Lys

Molecular architecture of the myelinated axon



Niemann, Berger and Suter, *NeuroMolecular Medicine*, 2006;8:217-24 (updated)

What have we learnt? **One can perturb the neuron/nerve in a multitude of ways = CMT**

Genetic contributions to inherited and apparently acquired neurologic dz

CMT: clinical & genetic aspects

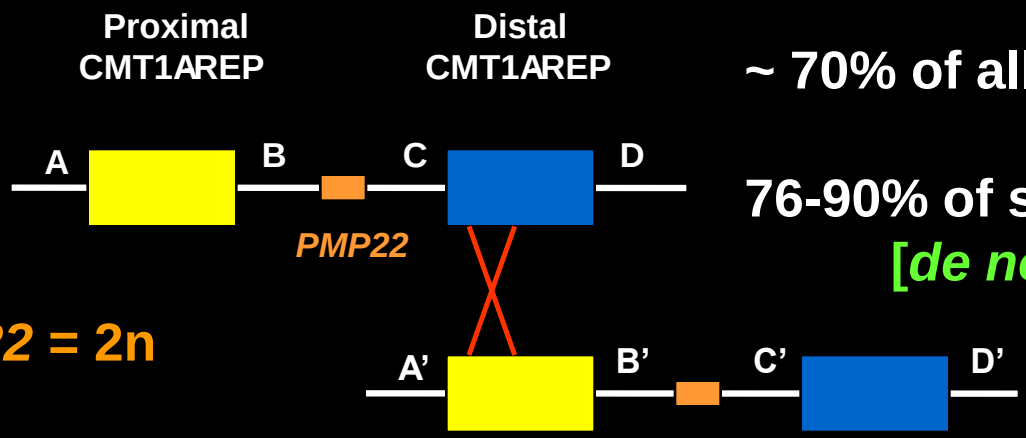
The CMT1A duplication
- a paradigm for CNV mutation & mechanisms for CNV formation

CMT mutational load
- gene load? locus load? or genomic load?
- SNP + CNV

Personal genome sequencing: CMT & DRD

The CMT1A duplication – a CNV paradigm

Raeymakers, Timmerman, *et al*. (1991) *Neuromuscular Disorders 1* :93-97
Lupski, *et al*. (1991) *Cell 66* :219-232; Lupski, *et al* (1992) *Nat Genet 1* :29-33
[duplication, gene dosage] ; Pentao, Liu, *et al* (1992) *Nat Genet 2* :292-300 [NAHR]

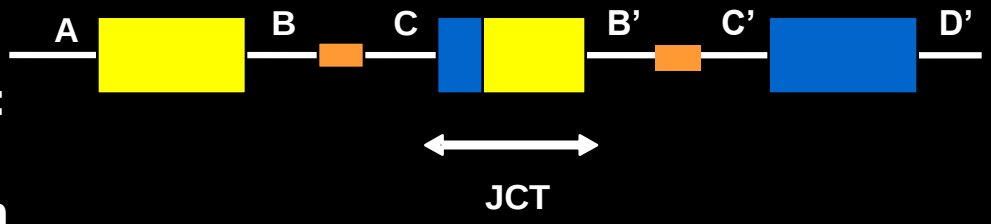


~ 70% of all CMT1 pts

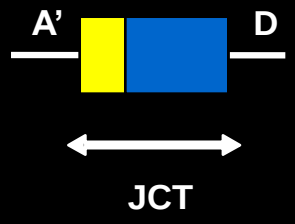
76-90% of sporadic CMT1
[*de novo* mutation]

NORMAL: *PMP22* = 2n

CMT1A: *PMP22* = 3n
HNPP: *PMP22* = 1n

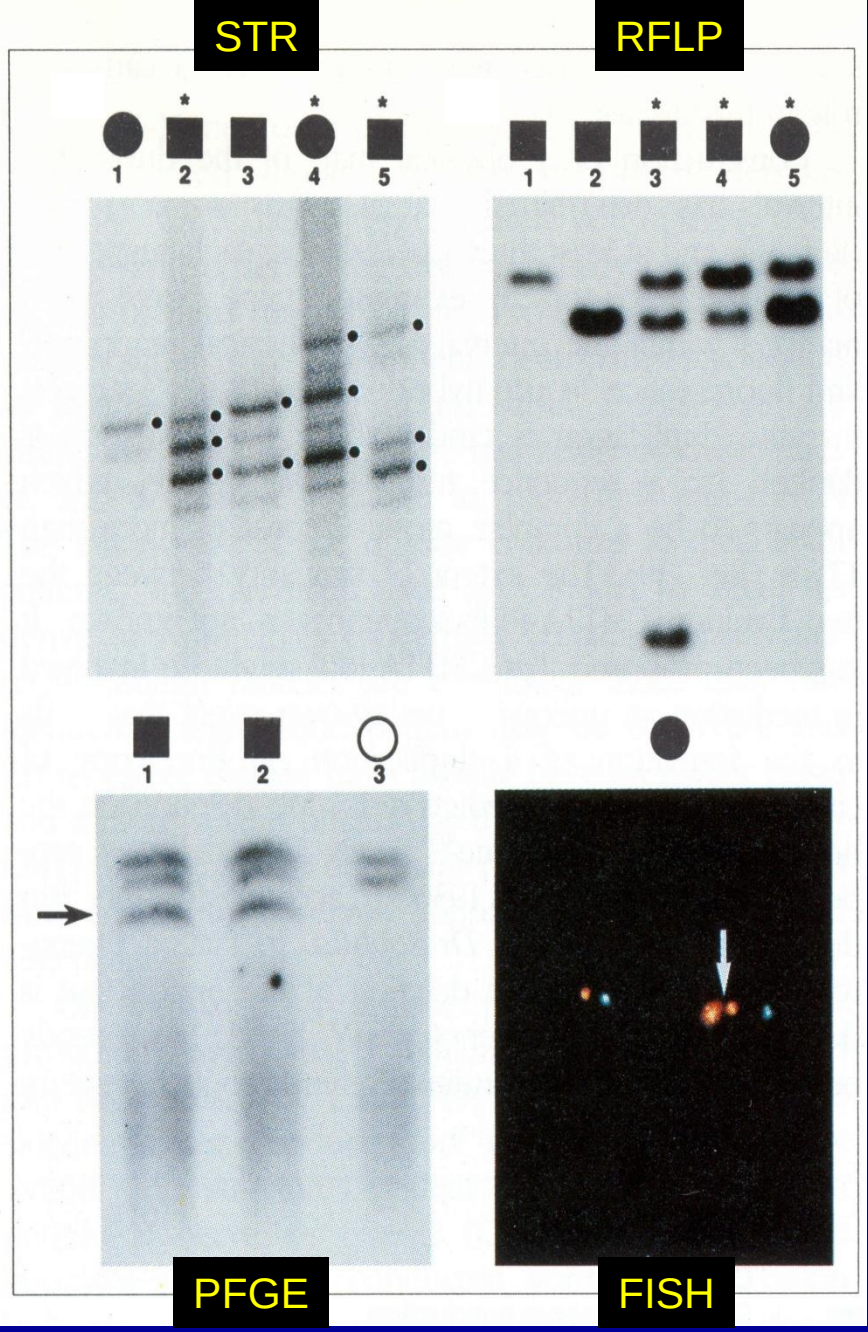


CMT1A DUPLICATION



HNPP DELETION

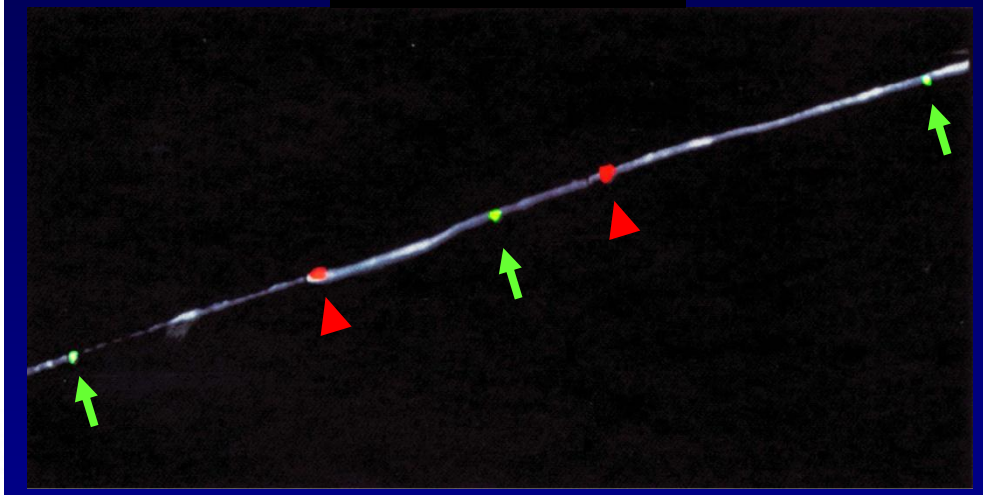
CNV dzs:
SCZ
Autism
Obesity



Patel and Lupski (1994) *TIG 10*:128-133

independent molecular methods reveal evidence for CMT1A duplication

fiber-FISH



Rautenstrauss (1997) *J Periph Nerv Sys 3*:1-4

What have we learned?

CNV associated with genomic disorders highlight:

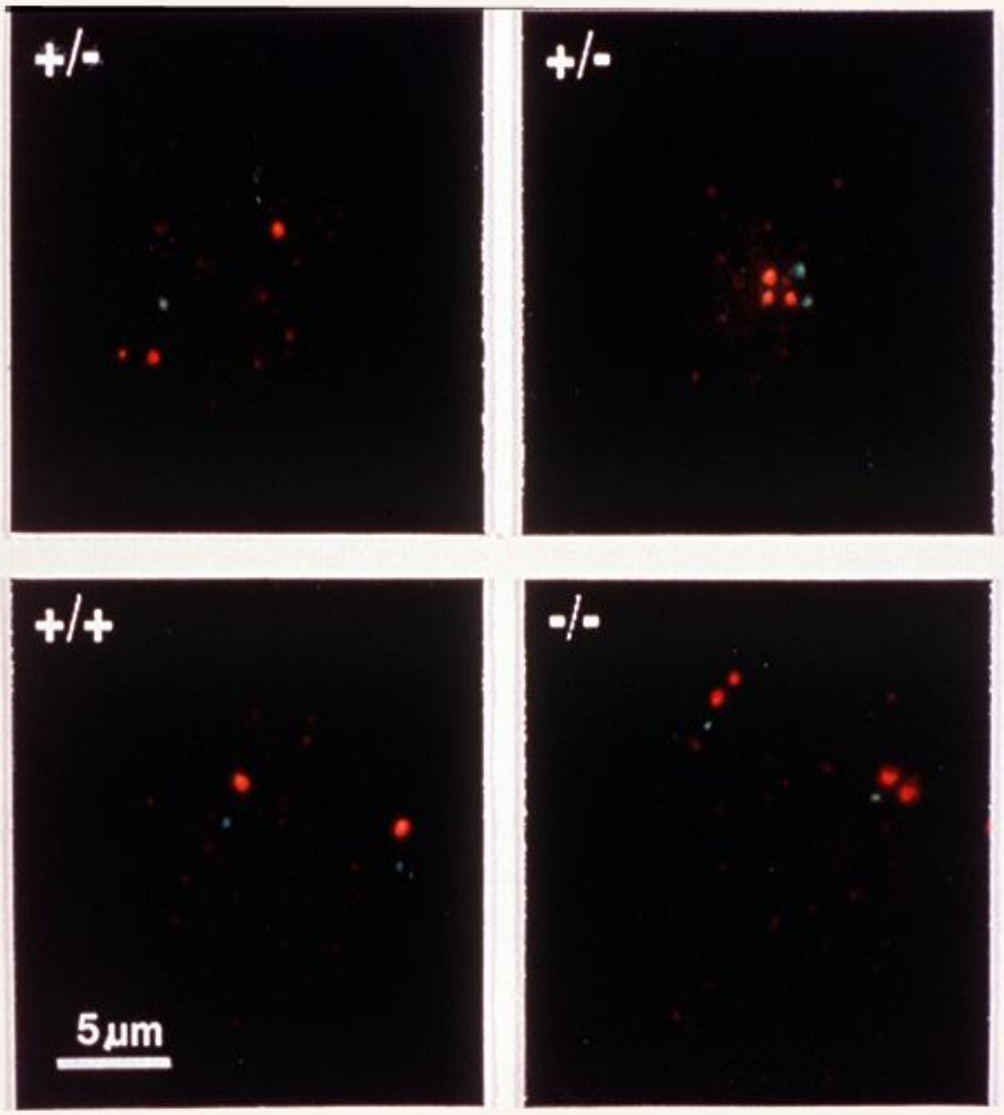
i) **Disease allele transmission; dup CNV = triallelic**

ii) **Gene dosage role in clinical traits**

Contrasting features of CMT1A and HNPP

	CMT1A	HNPP
Clinical	Symmetric, slowly progressive	***Asymmetric, episodic***
Antecedent	None	Motor nerve compression
Potential Early Signs	Mild delay in achieving motor milestones Idiopathic toe walking of childhood Absent deep tendon reflexes	None
Presentation	Distal muscle weakness and atrophy Dropped foot abnormal gait Foot deformity (*pes cavus, pes planus*)	Pressure palsies ***Focal neuropathy*** ***Carpal tunnel syndrome***
Electrophysiologic	Slow NCV	Conduction block
Neuropathology	Onion bulb	Tomacula
Molecular	Duplication	Deletion

HMZ dup gives severe disease: gene dosage!



Lupski *et al*. 1991
Cell 66:219-232

PHENOTYPIC VARIABILITY

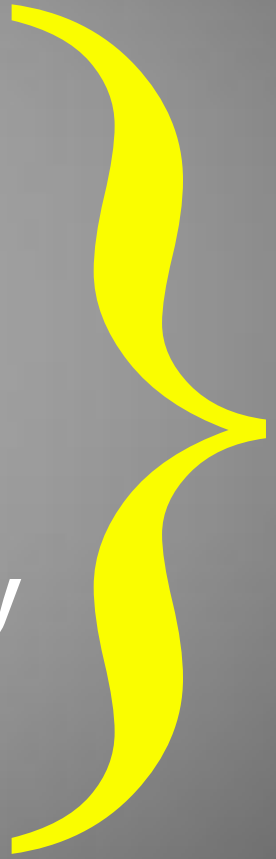
Charcot-Marie-Tooth type 1

Roussy-Levy syndrome

Dejerine-Sottas syndrome
(hmz > htz)

HMSN with calf hypertrophy

Scapuloperoneal atrophy
(Davidenkow syndrome)



CMT1A
duplication

Ann. Neurol. (1996) *39*:180-186

Deletions of Chromosome 17p11.2 in Multifocal Neuropathies

J. Tyson, BSc,* S. Malcolm, PhD,* P. K. Thomas, DSc,†‡ and A. E. Harding, FRCP†

- **24/51 patients with multifocal neuropathy have HNPP deletion**
- **7/19 (37%) index cases had no affected relatives**
- **Peripheral nerve lesion related to pressure in only 62% of cases**

Multifocal neuropathy genetic?

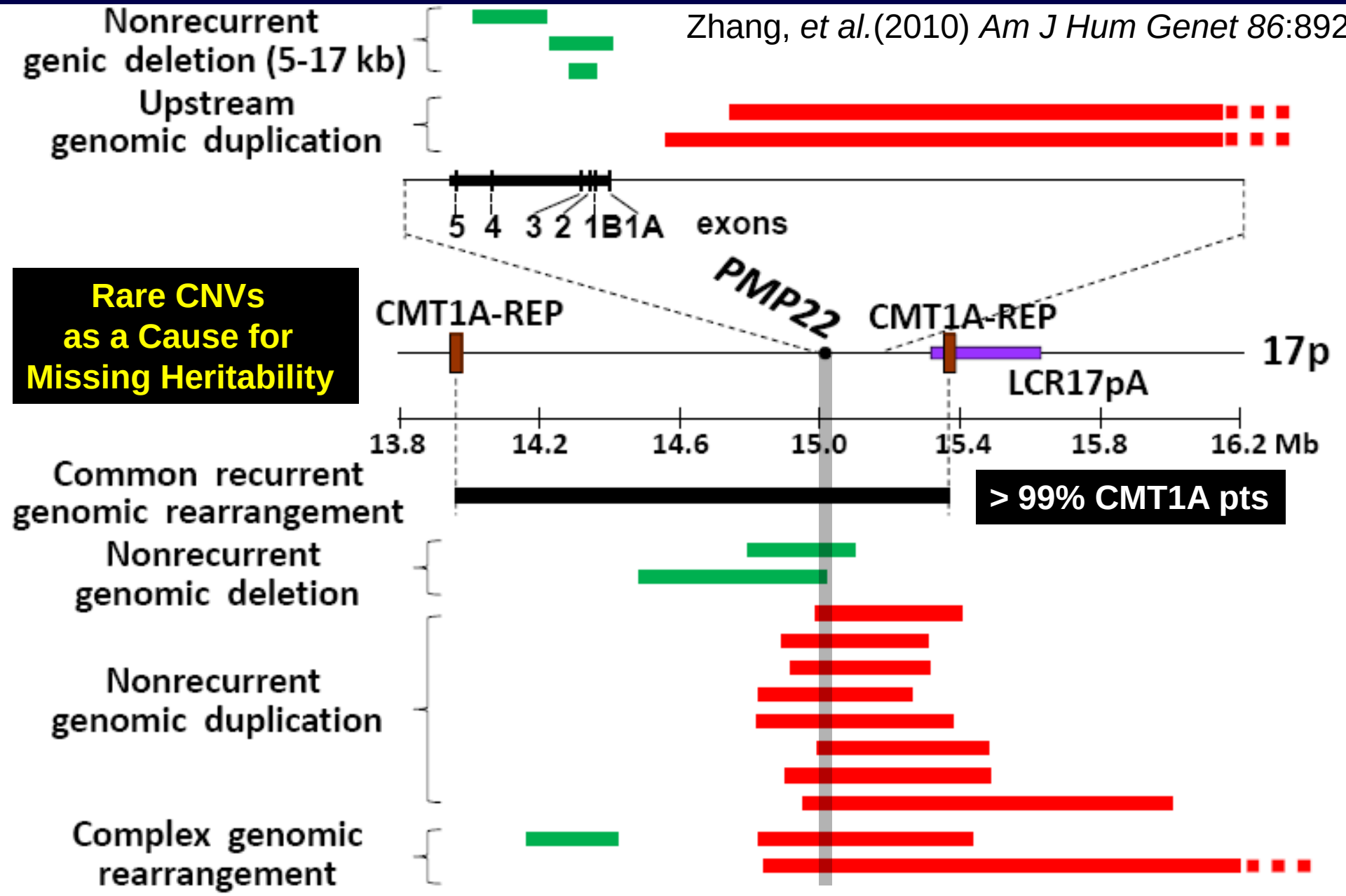
PMP22 CNV detected by abnormal MLPA for CMT1A duplication

Year	dup/del test	nml	dup	del
2007	4261	3472	549	194

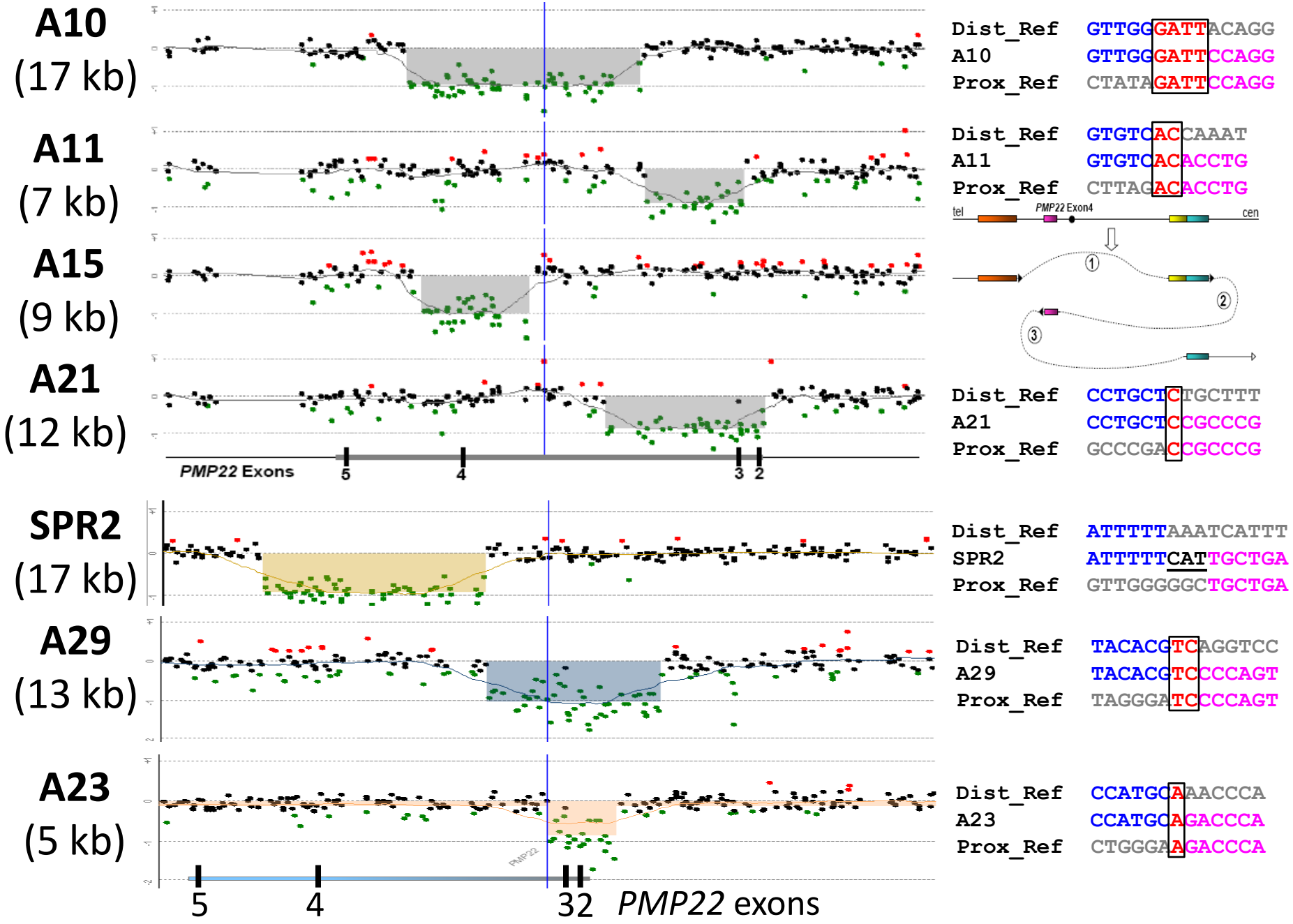
- MLPA unusual in 7 samples
- Frequency of detecting dup or del = (549+194)/4261 = **17.5%**
- Frequency of unusual MLPA = 7/(549+194) = **0.8%**
- Estimated NAHR at CMT1A/HNPP locus = 99.2%
- del/dup = 194/549 = ~.35 (NOT 2:1); ~80% HNPP undiagnosed!

Zhang *et al* (2009) *Nature Genetics 41*:849-853

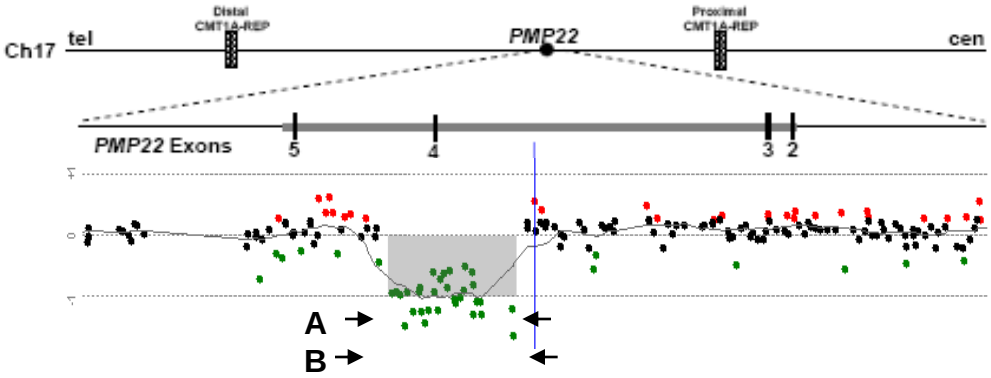
CNVs involving the coding or upstream regions of *PMP22* from patients with CMT1A or HNPP

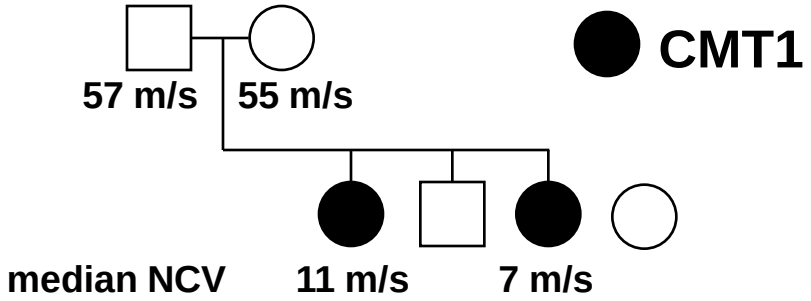


Exonic deletions of *PMP22*



FoSTeS caused complex deletion of *PMP22* exon 4



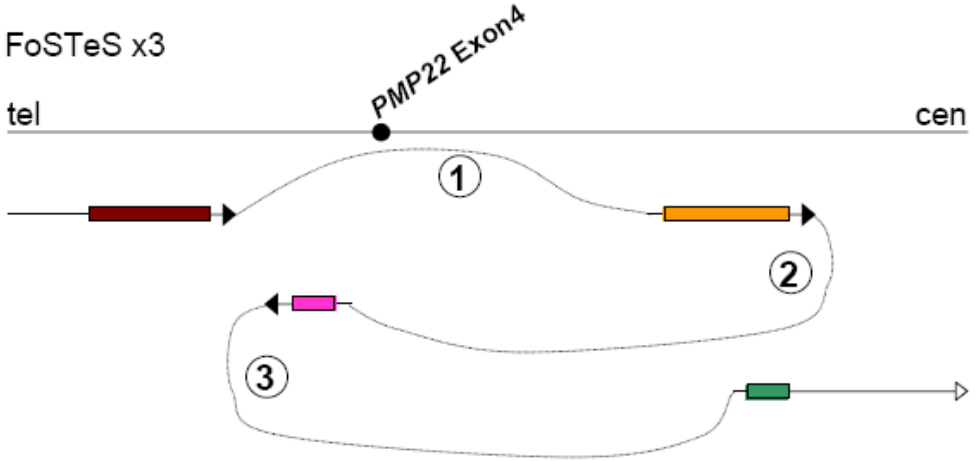


FoSTeS 1

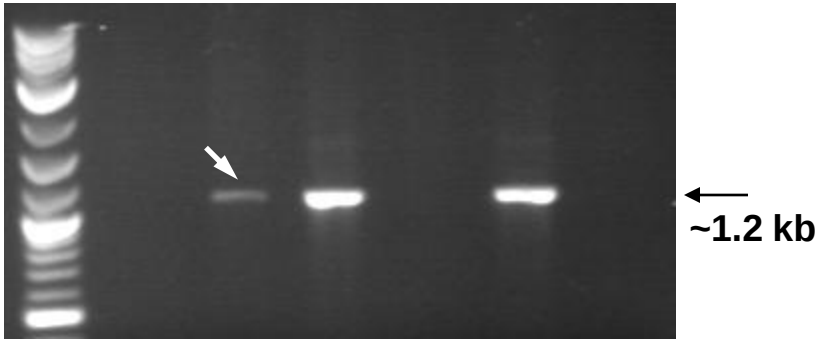
```
DistRef1+  TTGATGTTTTCCAGTCTAGTGCAACCCAACCCAACAGATCTTGTCAAGGAATAGATGGCTATAGGTTCTG
A15_1+     TTGATGTTTTCCAGTCTAGTGCAACCCAACCCAACAAAGAGAAAACAGCTAAGTATAAAATTGAAAAGCC
ProxRef1+  TCGCATCATTAACAAAATTAAATTACAGACAGAACAAAGAGAAAACAGCTAAGTATAAAATTGAAAAGCC
```

FoSTeS 2 & 3

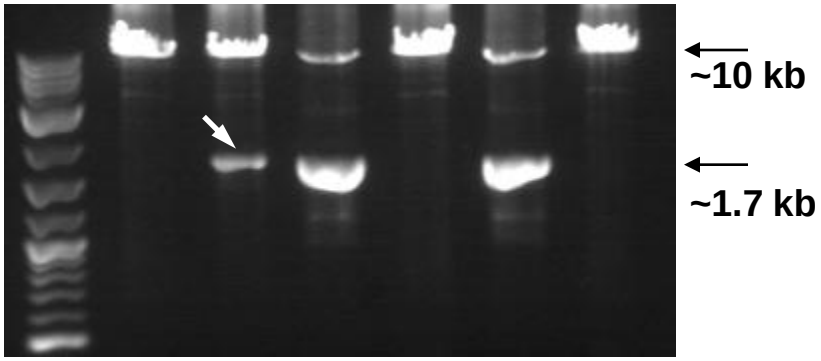
```
DistRef2+  TCGTAAAAGGTGCCCAACCTCACTAGCAACCAAGGAAATGCAAGAGAAACCCATGAGGAGGGTGACACCA
A15_23+    TCGTAAAAGGTGCCCAACCTATAGCCATCTATTCCTTGACAAGGTGCCCAACCTCACTAGCAACCAAGGA
ProxRef23- CATTCTTATTTTCAGAACCTATAGCCATCTATTCCTTGACAAGATCTGTTGGGTTGGGTTGCACTAGACT
DistRef3+  TTAGCAAAGGAGAAATATGAACAGCCAATAAACATCGTAAAAGGTGCCCAACCTCACTAGCAACCAAGGA
```



Mosaic complex rearrangement in mother suggests mitotic event consistent with the MMBIR/FoSTeS model



standard PCR; primers A



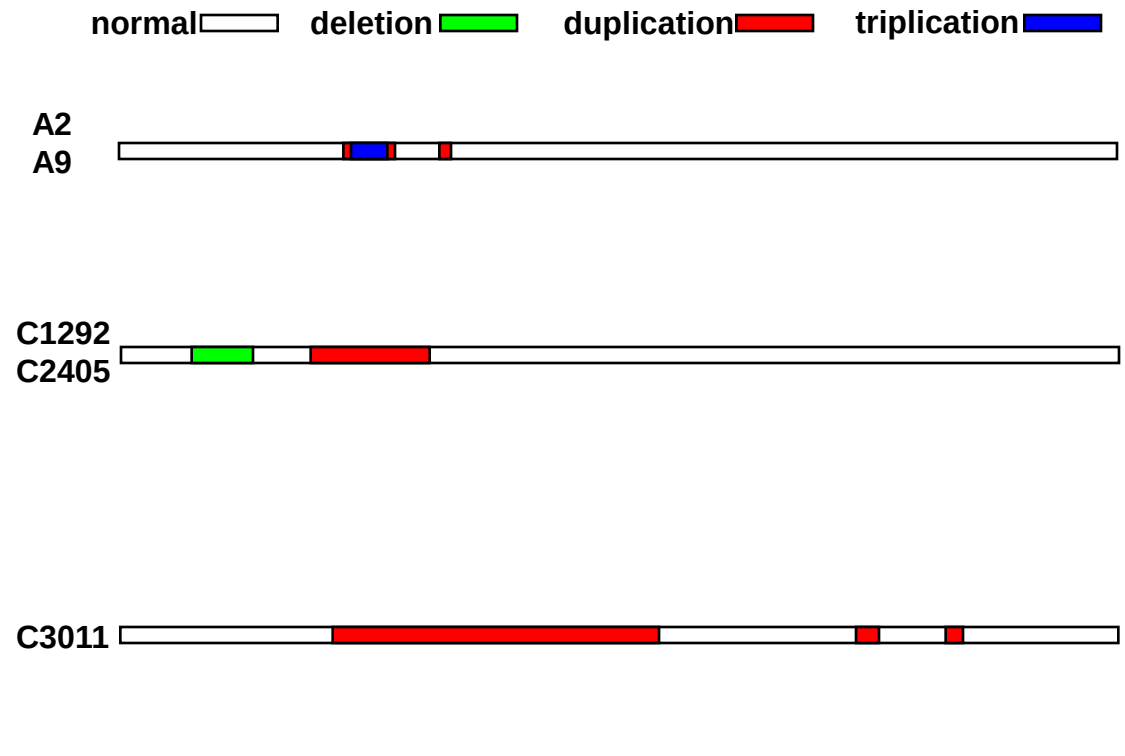
long-range PCR; primers B

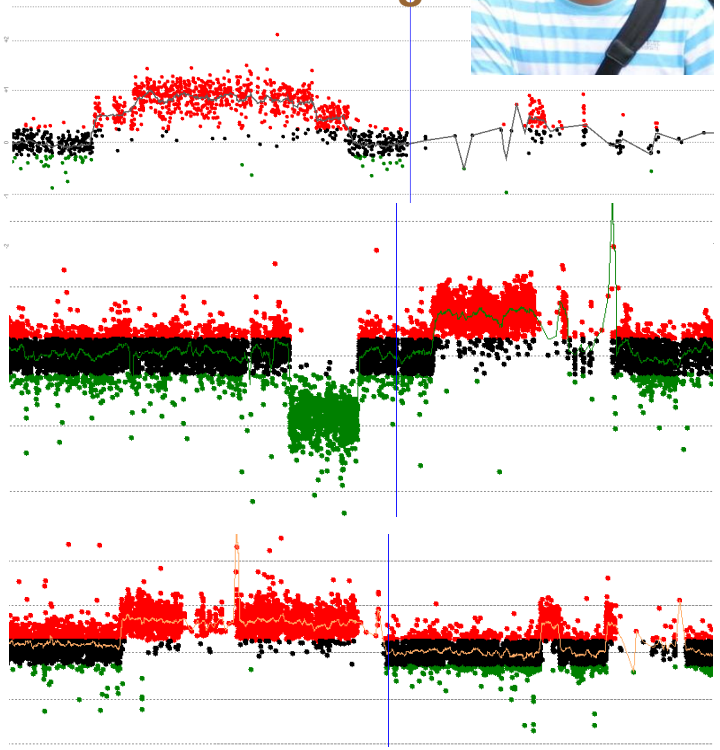
mosaic deletion in mother

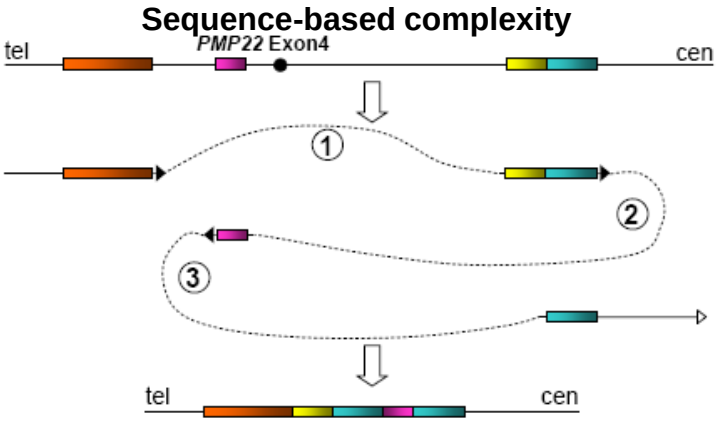
Zhang *et al*. (2009) *Nat Genet 41:*849-853

Complex CMT1A Rearrangements

Feng Zhang







Types of rare CNVs observed at the CMT1A *PMP22* locus

Nonrecurrent genomic duplications and deletions

Complex rearrangements

Exonic deletions

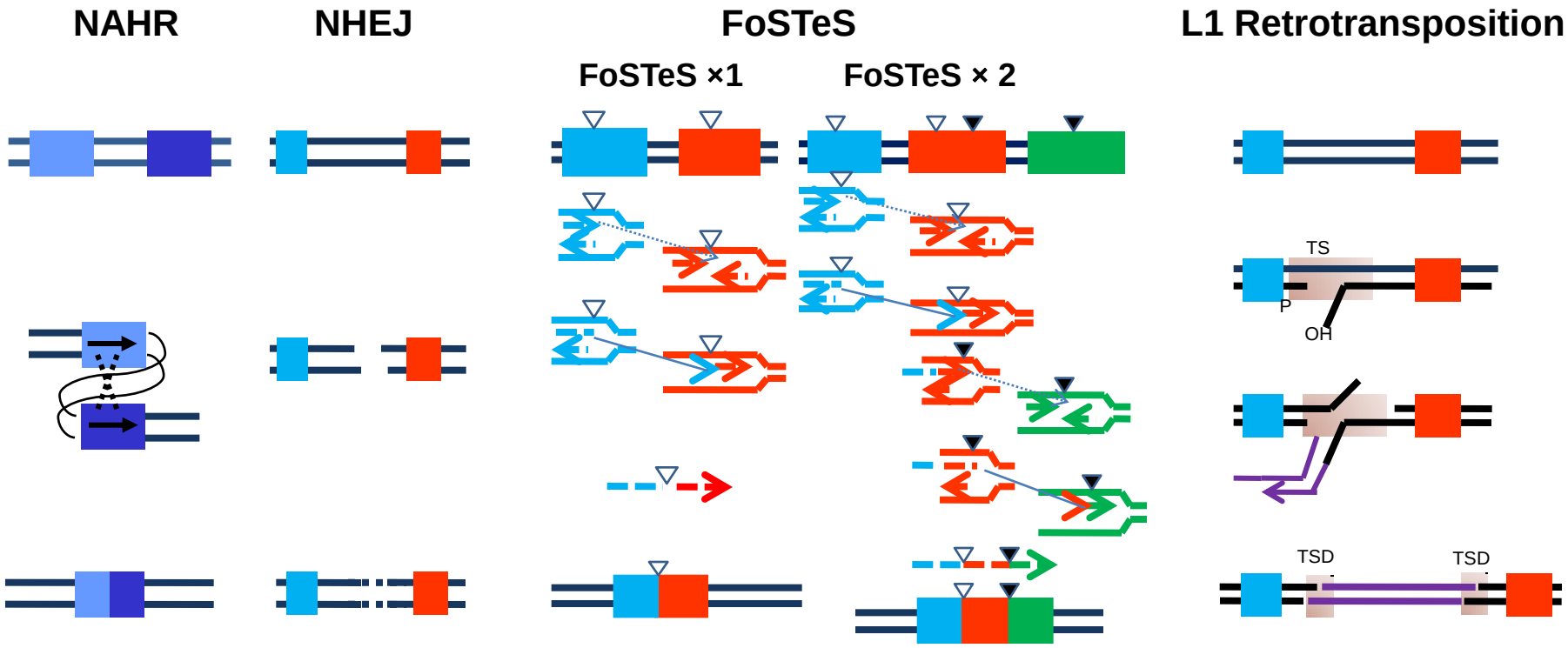
Upstream sequence duplications

A patient with multiple *de novo* rearrangements

Mechanisms for genomic disorder associatred human genomic rearrangements

MMBIR: *microhomology-mediated, break induced replication*

MEI *– mobile element insertion*



Zhang, Gu, Hurles, Lupski (2009) *Ann Rev Genomics and Hum Genet 19*:451-481

Cell 131:1235-1247, December 26, 2007

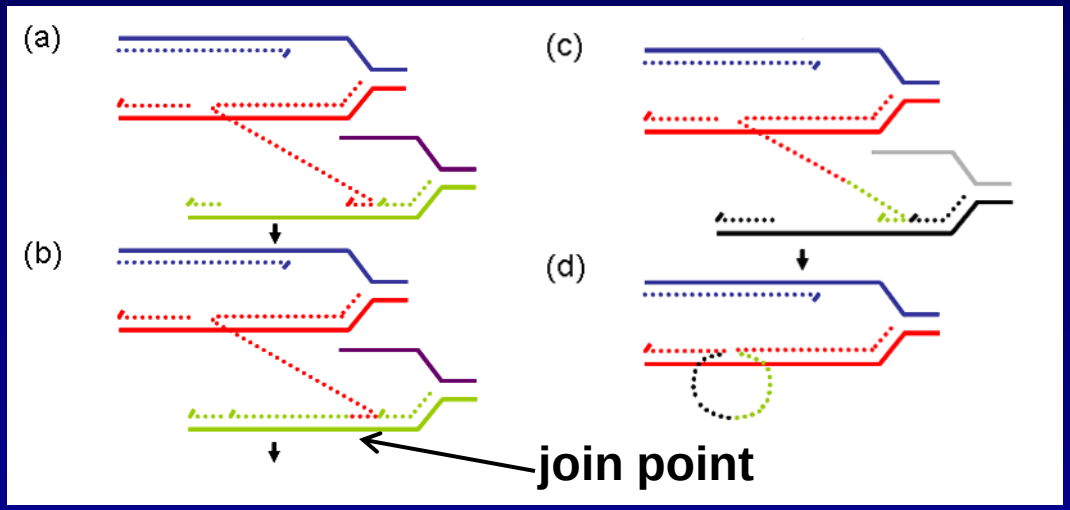
Cell

A DNA Replication Mechanism for Generating Nonrecurrent Rearrangements Associated with Genomic Disorders

Jennifer A. Lee,[1] Claudia M.B. Carvalho,[1] and James R. Lupski[1,2,3,*]

Jenny Lee

Claudia Carvalho

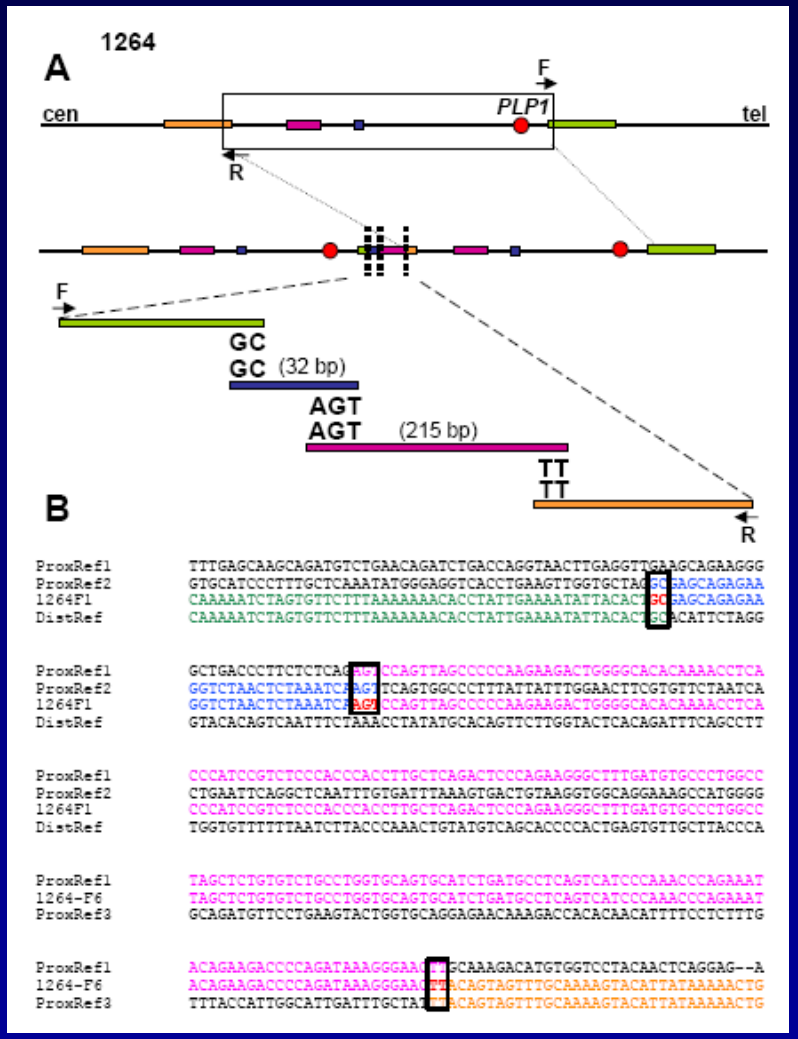


- Studied Pelizeaus-Merzbacher Dz
- CNS dysmyelinating disorder
- ~ 70% due to different sized (i.e. non-recurrent) *PLP1* dup

DNA replication mechanism:
Fork Stalling
Template Switching,
FoSTeS

1) Long distance template switching (120-550 Kb)
2) Tethered to original fork
3) Priming of DNA replication via microhomology
4) Template driven juxtaposition of discreet genomic segments from different locations

DNA replication model for genomic rearrangements



1264

Lee et al. (2007) *Cell 131*:1235-1247

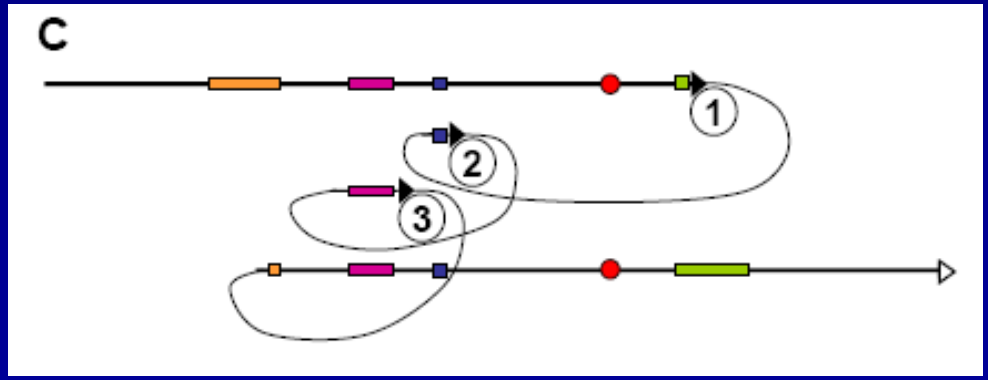
Fork Stalling and Template Switching

FoSTeS

Mechanism

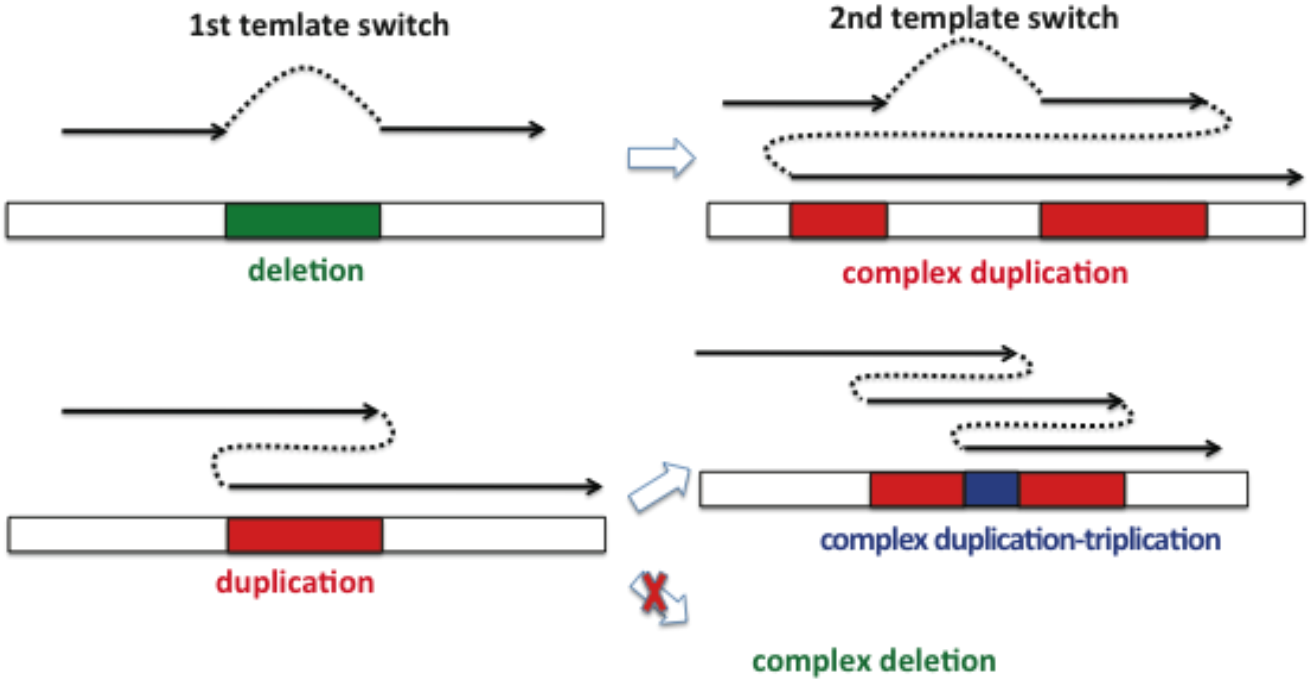
- Microhomology mediated joining

- Template driven juxtaposition of DNA sequences separated by large genomic distances



FoSTeS x 3

FoSTeS/MMBIR favors gain (DUP, TRP, etc.) over loss of genomic material



Pengfei Liu

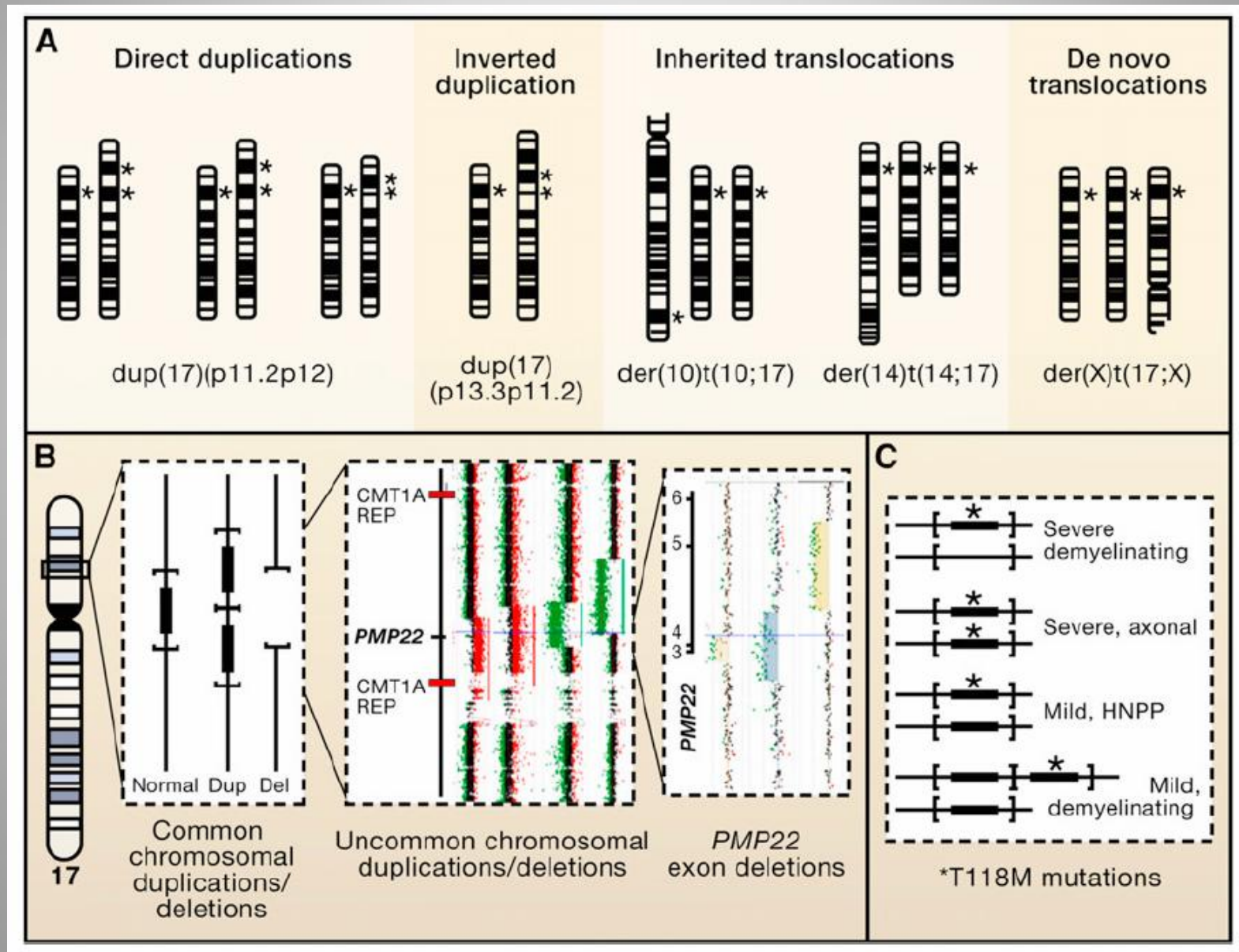
A 2nd template switch can erase the deletion generated in the first step.

However, a duplication in the first step can never be erased.

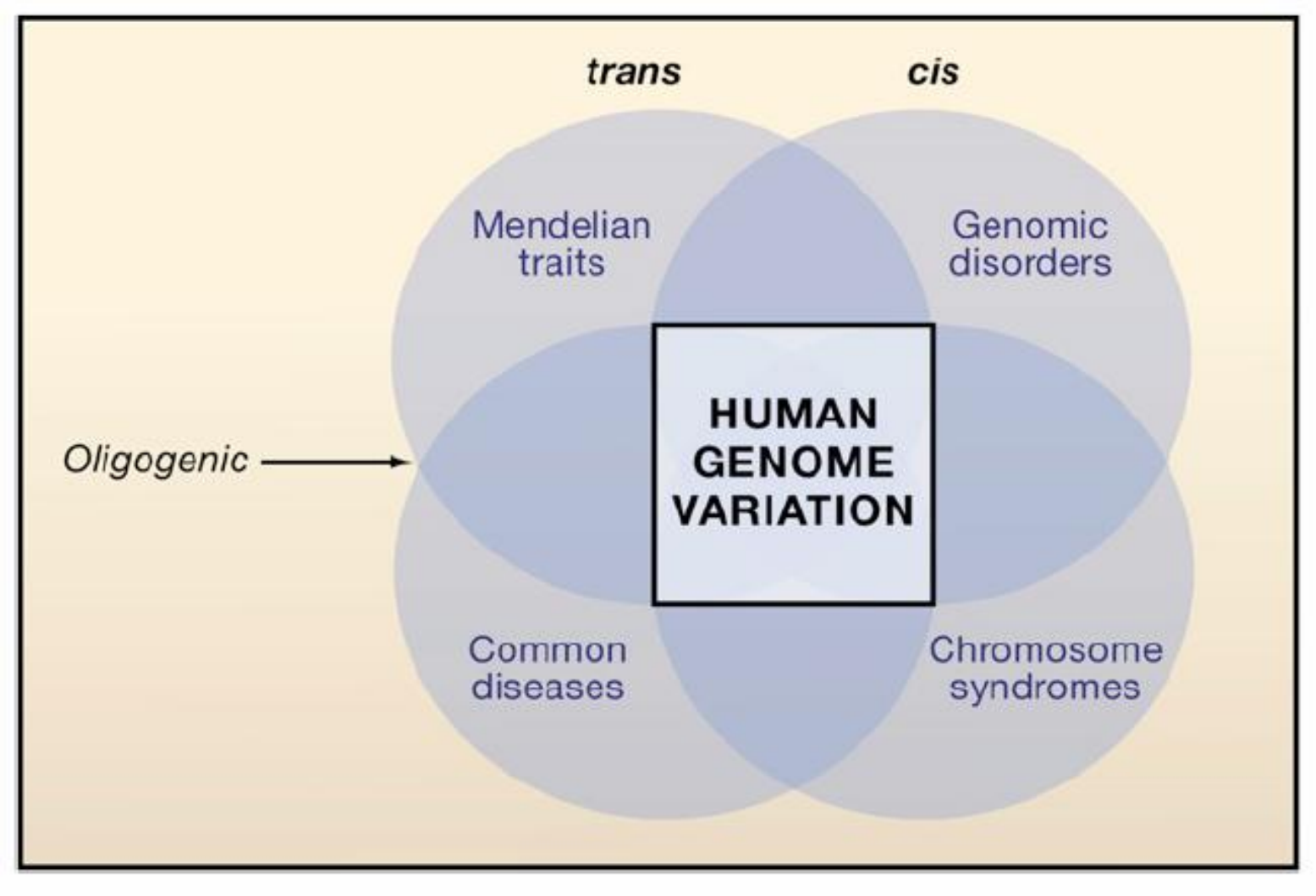
Replicative mechanism important to evolution: i) gene duplication/triplication ii) exon shuffling

Liu, *et al.* (2011) *Am. J. Hum. Genet. 89*: 580-588

Bridging the gap between chromosomal syndromes and Mendelian disorders



A continuum for the genetics & genomics of human disease



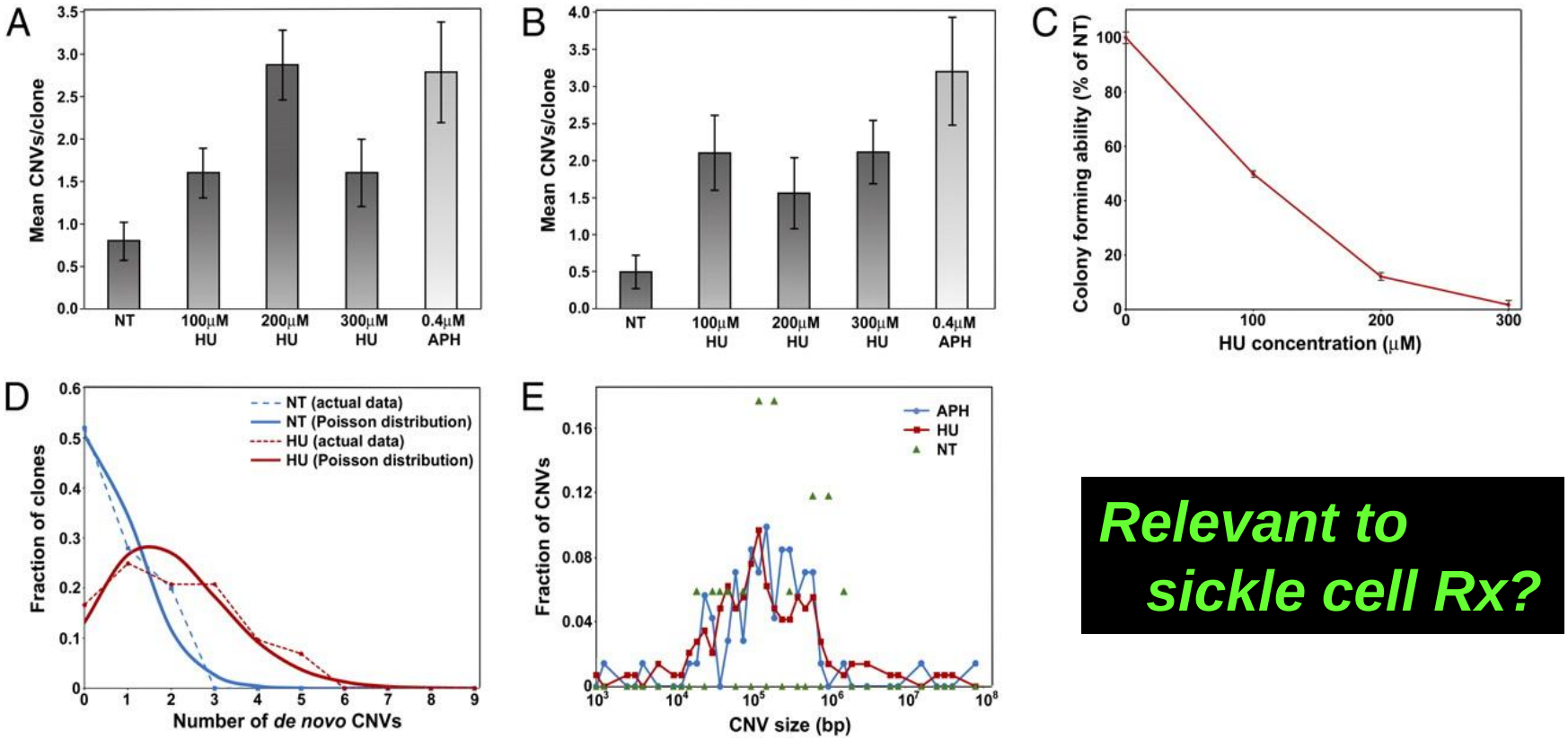
J.R. Lupski, J.W. Belmont, E. Boerwinkle, R.A. Gibbs (2011) *Cell 147*: 32-43

Hydroxyurea induces de novo copy number variants in human cells

Martin F. Arlt[a], Alev Cagla Ozdemir[a], Shanda R. Birkeland[b], Thomas E. Wilson[a,b], and Thomas W. Glover[a,1]

Arlt M F et al. *PNAS* 2011;*108*:17360-17365

HU induces de novo CNVs in normal human fibroblasts.



Relevant to sickle cell Rx?

Genetic contributions to inherited and apparently acquired neurologic

CMT: clinical & genetic aspects

The CMT1A duplication
- a paradigm for CNV mutation

CMT mutational load
- gene load? locus load? or genomic load?
- SNP + CNV

Personal genome sequencing: CMT

Allelic series with combined CNV & SNP

T118M *PMP22* Mutation Causes Partial Loss of Function and HNPP-like Neuropathy

Michael E. Shy, MD,[1] Mena T. Scavina, DO,[2] Alisa Clark, RN, MSN,[2] Karen M. Krajewski, MS,[1] Jun Li, MD, PhD,[1] John Kamholz, MD, PhD,[1] Edwin Kolodny, MD,[3] Kinga Szigeti, MD,[4] Richard A. Fischer, MD,[2] Gulam Mustafa Saifi, PhD,[4] Steven S. Scherer, MD, PhD,[5] and James R. Lupski, MD, PhD[4,6]

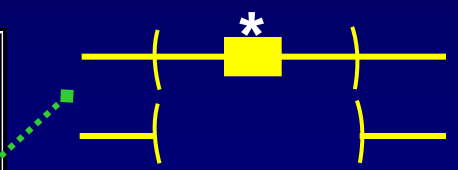
Ann Neurol 2006;59:358–364

Shy, *et al* (2006)
Annals of Neurology
***59:* 358-364**

Genotype-Phenotype Correlations

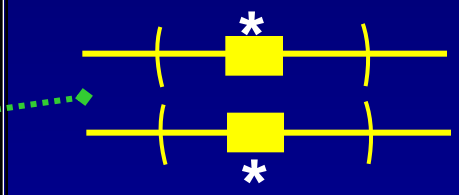
# of T118M alleles	# of wt alleles	Phenotype	Genotype
1	0	Severe, demyelinating	T118M/-
2	0	Severe, axonal	T118M/T118M
1	1	Mild, HNPP	T118M/+
1	2	Mild, demyelinating	T118M+/+

HNPP = hereditary neuropathy with liability to pressure palsies

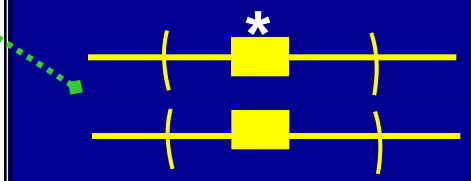


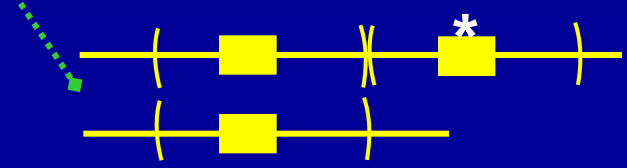
SNP + CNV

Roa, *et al* (1998) Nat Genet



SNP X2

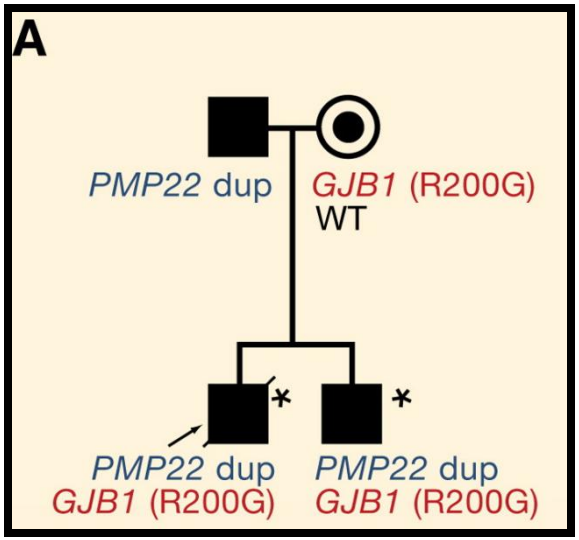


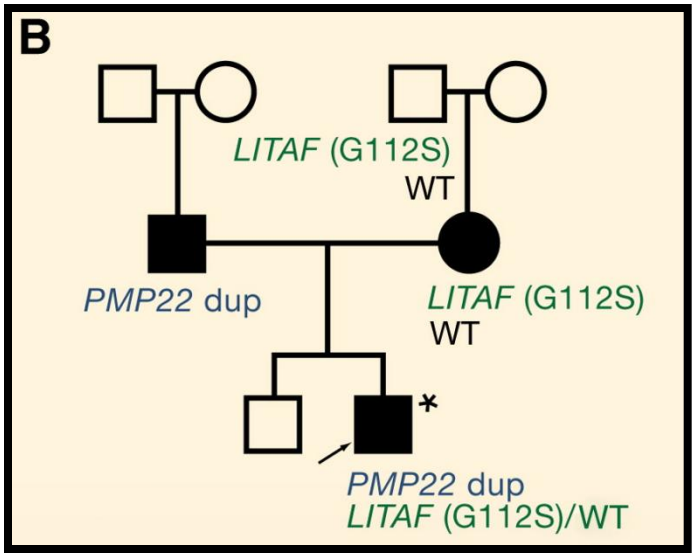


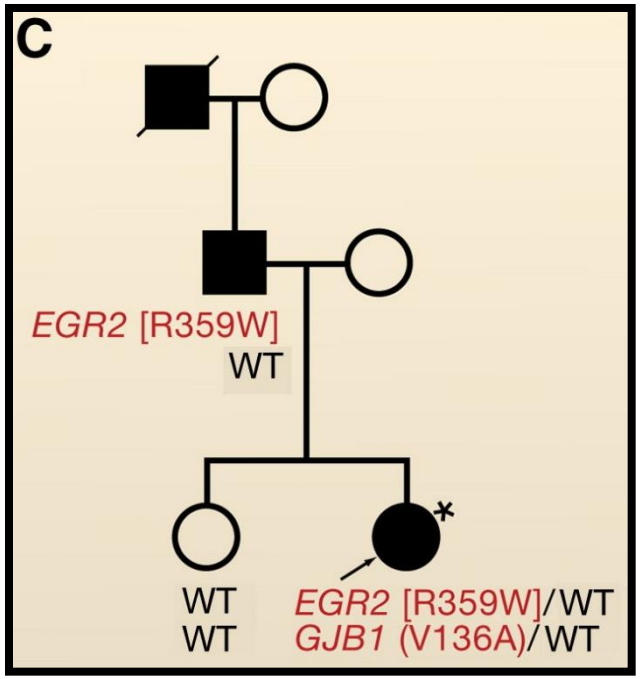
Genotypes: *PMP22* deletion ; **T118M** *

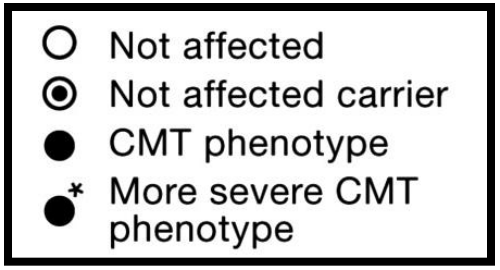
***PMP22* duplication**

Patients with mutation of two CMT genes

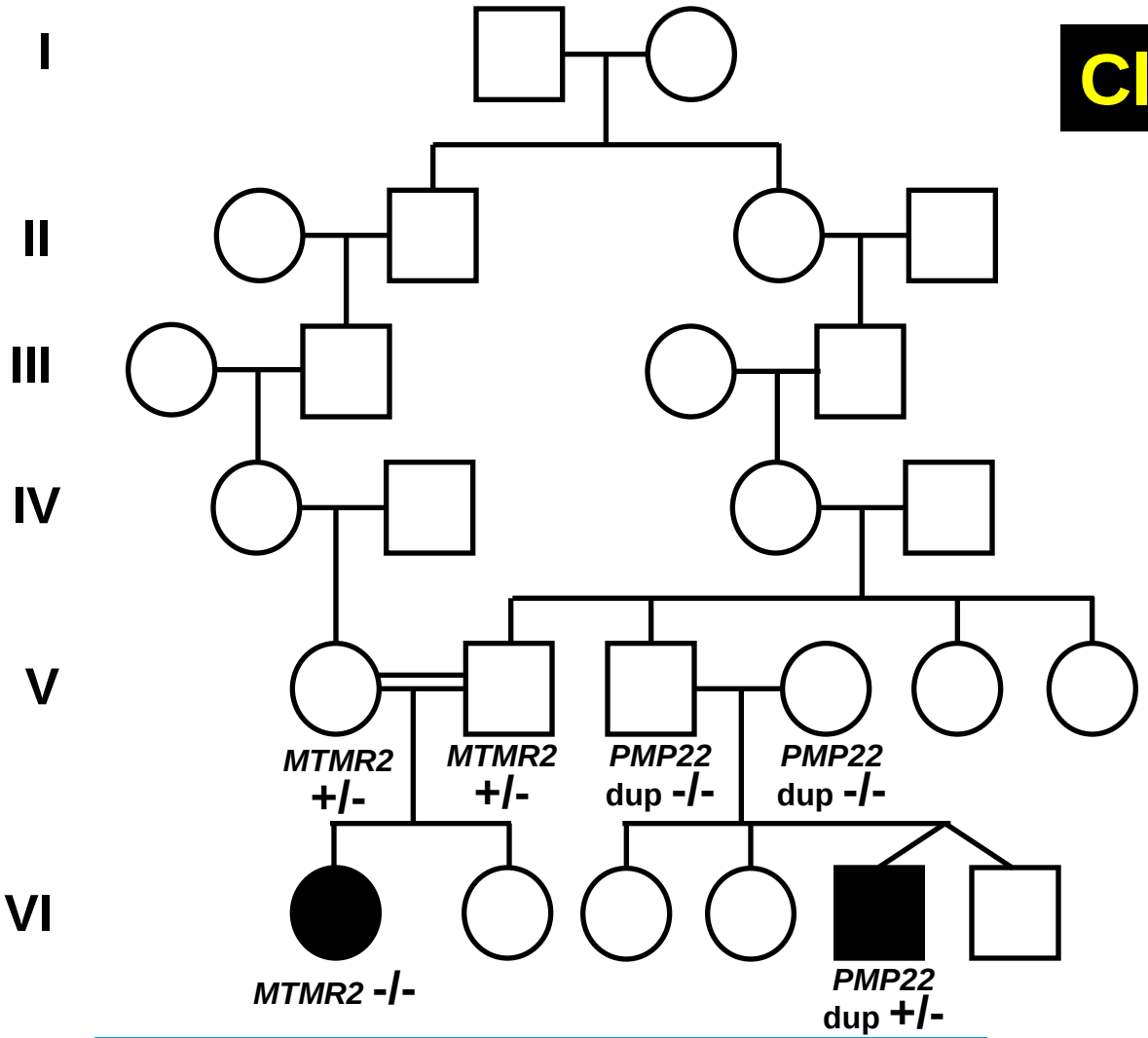








1 family, 1 gene, 2 genetic (AD & AR) forms for CMT



Clan Genomics!

- Richard Gibbs

The most important thing for individuals regarding their PERSONAL GENOME is what their nearest relatives gave them & *de novo* events.

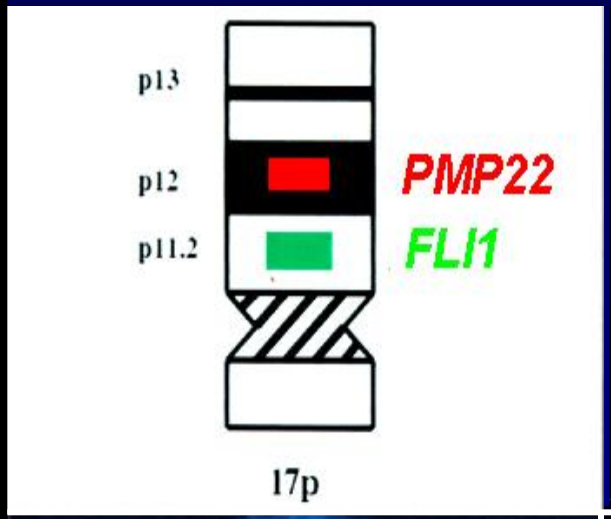
The population from which you come is NOT that relevant!

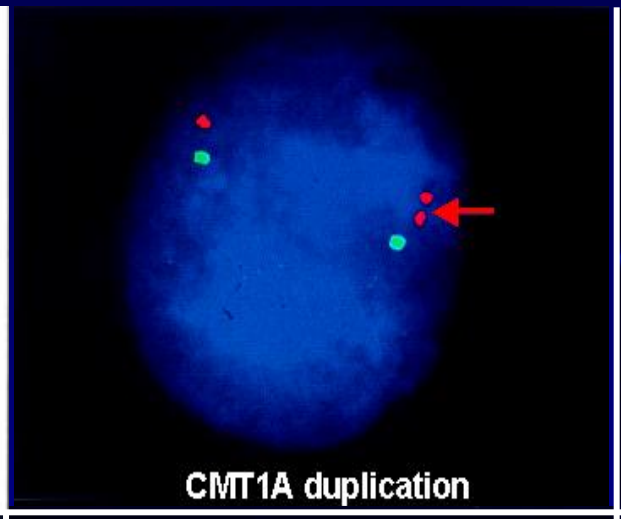
Statistical and population geneticist; No offense intended.

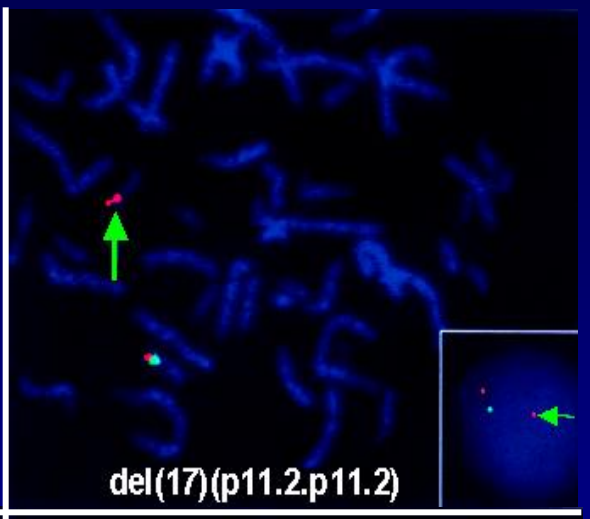
HMZ *MTMR2* mutation & *de novo PMP22* dup in same family !

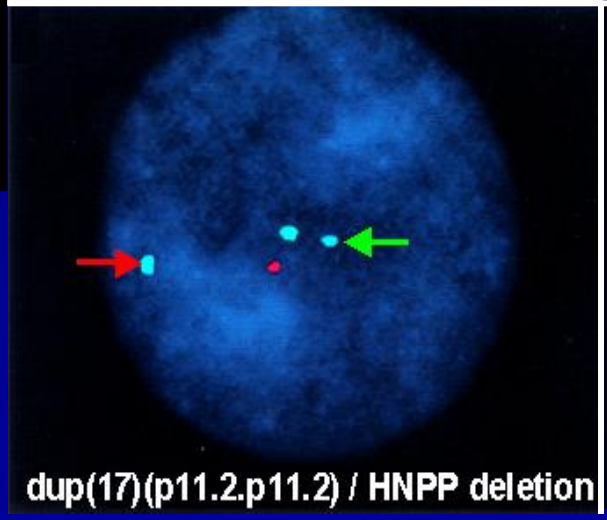
Verny C, *et al* (2004) *Neurology 63*; 1527-1529.

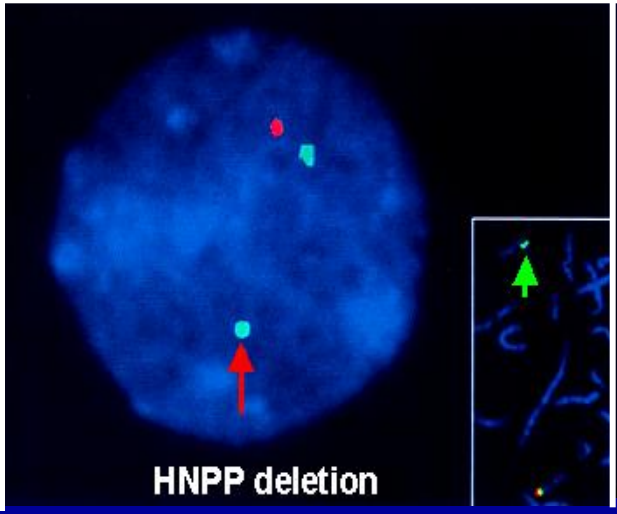
DNA rearrangements in 17p

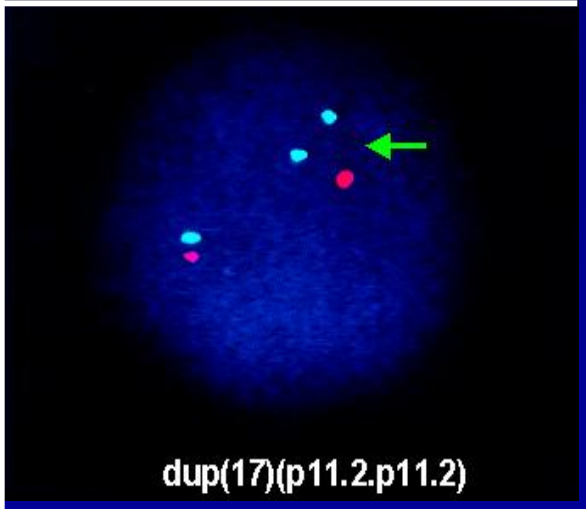










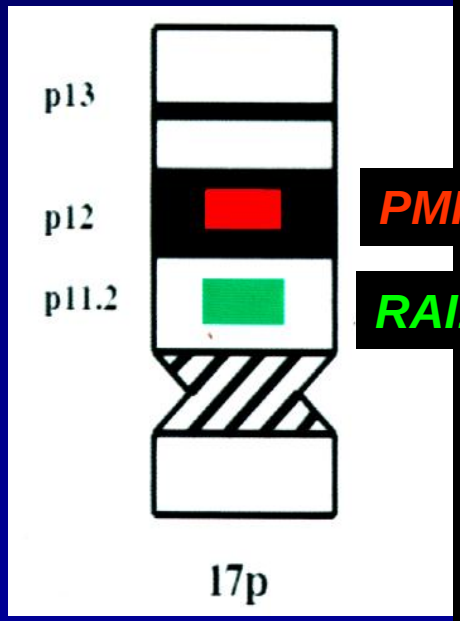


Inherited HNPP deletion segregating with carpal tunnel; de novo PTLS duplication!

DNA REARRANGEMENTS ON BOTH Ch17

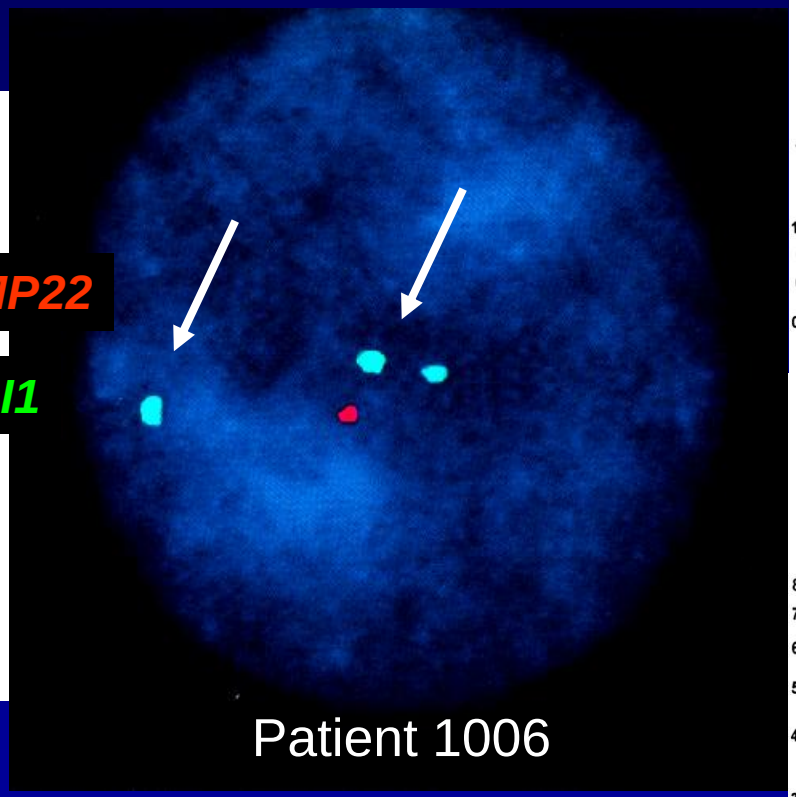
mildly delayed individual (PTLS) + neuropathy = complex trait?

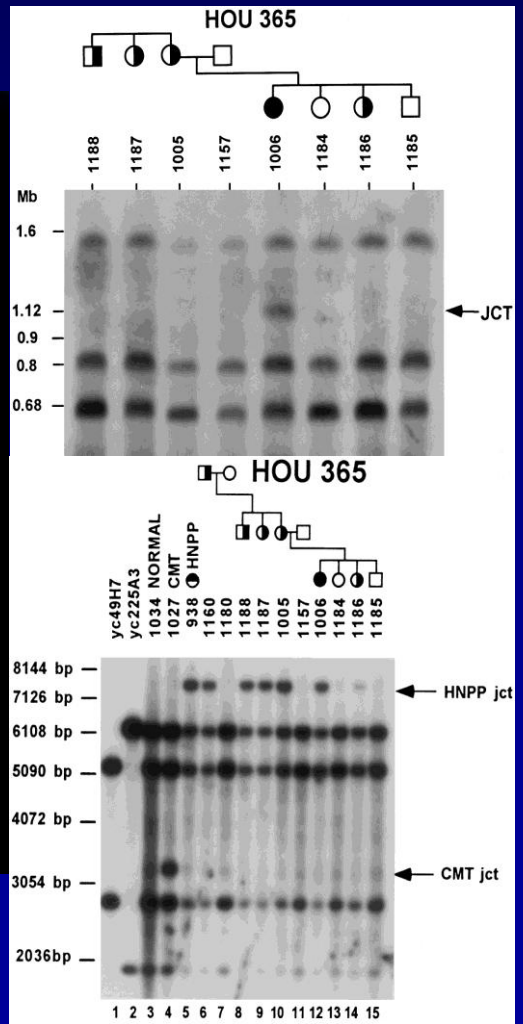
***de novo* + inherited alleles!**



PMP22

RAI1





Potocki et al. (1999)

***Am. J. Hum. Genet. 64*:471-478**

Genetic contributions to inherited and apparently acquired neuropathy

CMT: clinical & genetic aspects

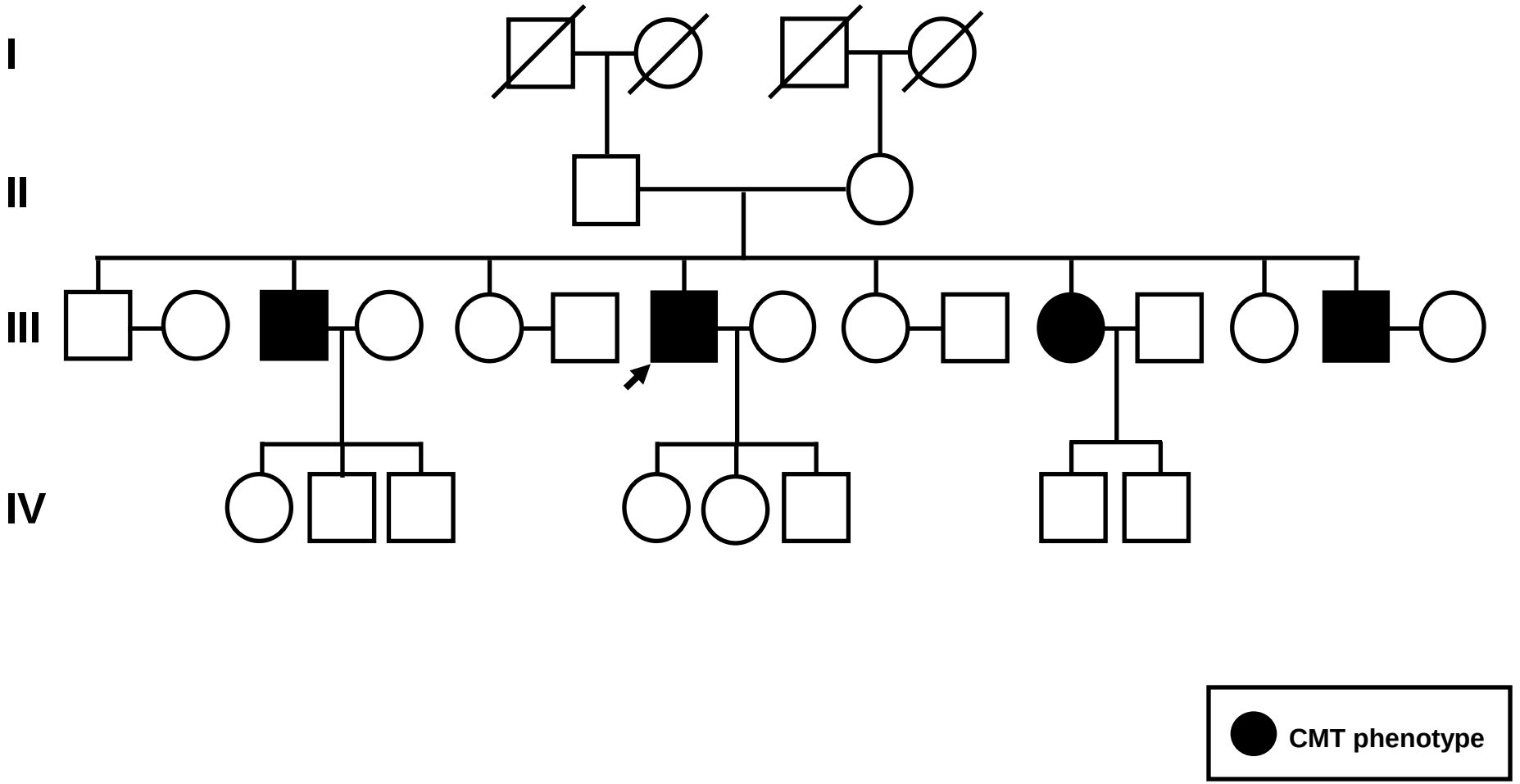
The CMT1A duplication
- a paradigm for CNV mutation

CMT mutational load
- gene load? locus load? or genomic load?
- SNP + CNV

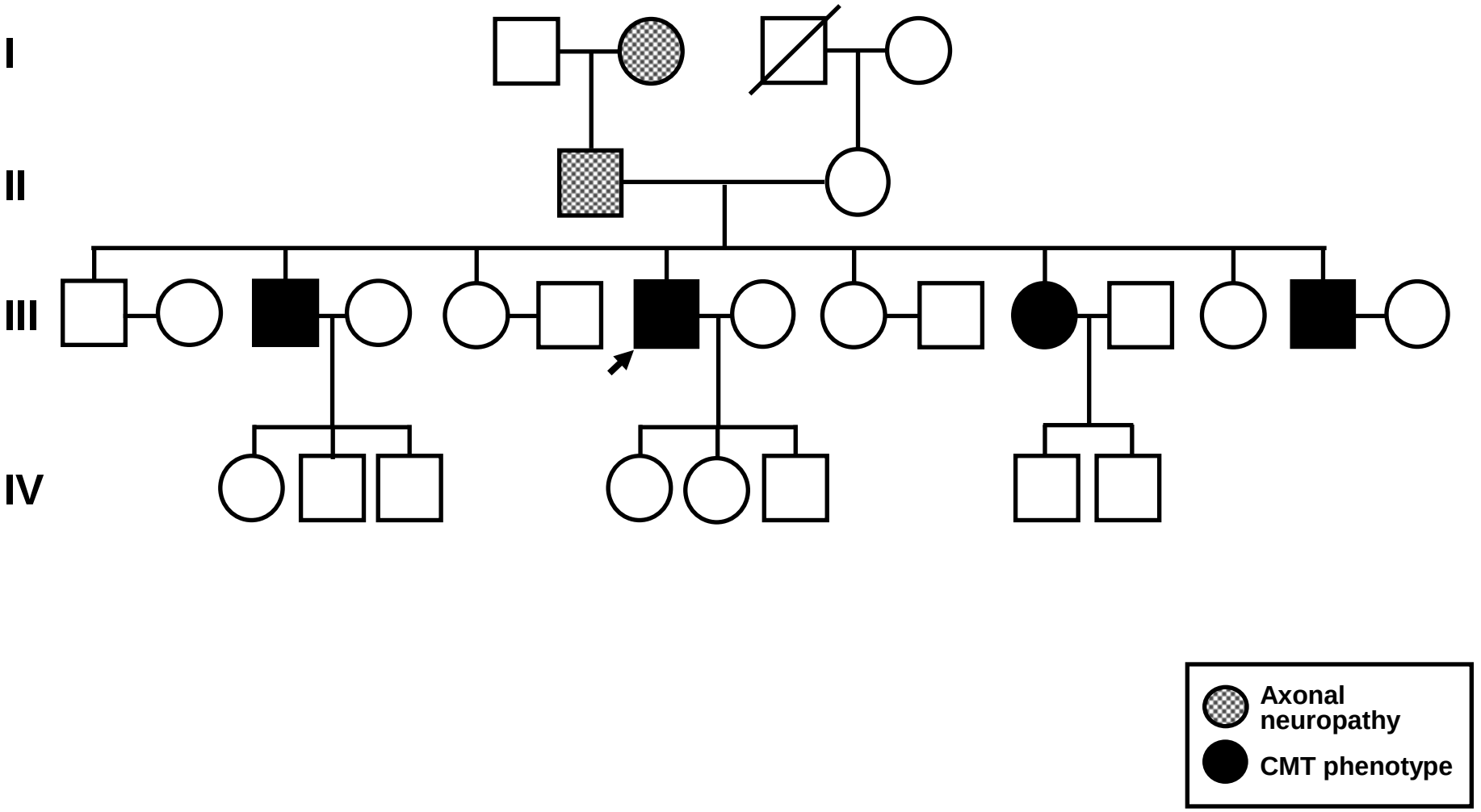
Personal genome sequencing:

CMT & the Lupski clan

Family: HOU37



Family: HOU37



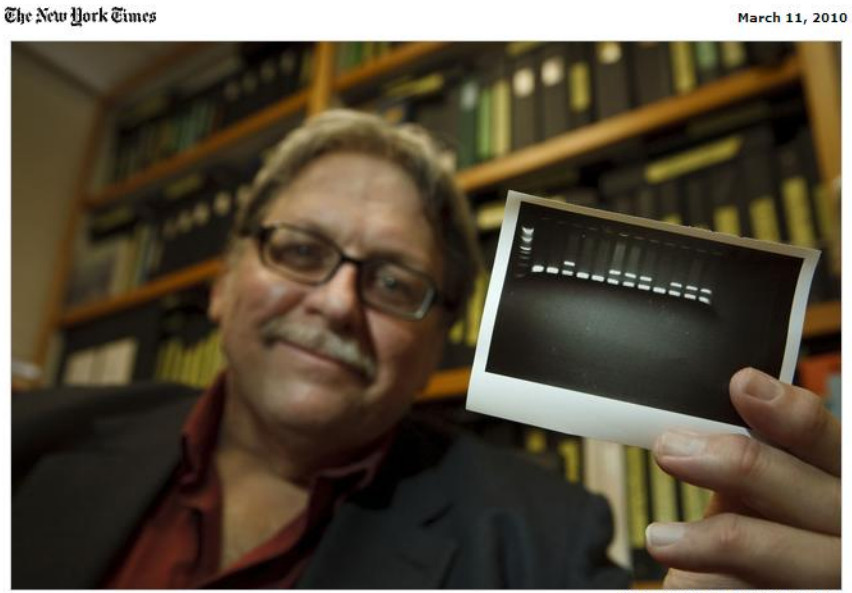
Something hard:

Finding disease alleles in a recessive disorder where the locus is unknown!

Subject: James R. Lupski

Disorder: (recessive) Charcot Marie Tooth (CMT) Syndrome

Approach: SOLiD WGS, family and functional follow-up



The New York Times March 11, 2010

Michael Stravato for The New York Times

TWITTER

SIGN IN TO RECOMMEND

Dr. James R. Lupski, a medical geneticist with a nerve disease, had his whole genome decoded.

SOLiD – Sequencing by Oligonucleotide Ligation and Detection (two base encoding method)

89.6 Gb mappable sequence; Average depth of coverage 29.6X

Human Genome Seq. Ctr.

HGSC
HUMAN GENOME SEQUENCING CENTER

BCM
Baylor College of Medicine

Richard A. Gibbs, Director

Comparison of complete Human Genomes

Individual	Ploidy	Technology	Average Depth	Total SNPs (M)	Known SNPs (M)	Novel SNPs (M)
Venter	2n	Sanger	7.5x	3.21	2.80	0.74
Watson	2n	Roche 454	7.4x	3.32	2.71	0.61
Chinese (YH)	2n	Illumina GA	36x	3.07	2.67	0.39
African (NA18507)*	2n	Illumina GA	40.6x	3.61	2.72	0.88
African (NA18507)*	2n	AB SOLiD	17.9x	3.86	3.13	0.73
Korean (SJK)	2n	Illumina GA	28.95x	3.43	3.01	0.42
Korean (AK1)	2n	Illumina GA	27.8x	3.45	2.88	0.57
Neuropathy subject	2n	AB SOLiD	29.6x	3.42	2.85	0.56

54 years from the Watson-Crick model of DNA to the J. D. Watson personal genome

Wheeler, *et al.* (2008)
***Nature* 452: 872-876**

Explaining clinical genetic implications of his personal genome to J.D. Watson

1) First personal genome by NGS
2) Tremendous variation!
3) Millions of bases, no match to ref
4) Both SNV & CNV vary a lot!
5) Allele frequency spectrum of CNV reveals smaller more freq
6) Challenging to interpret

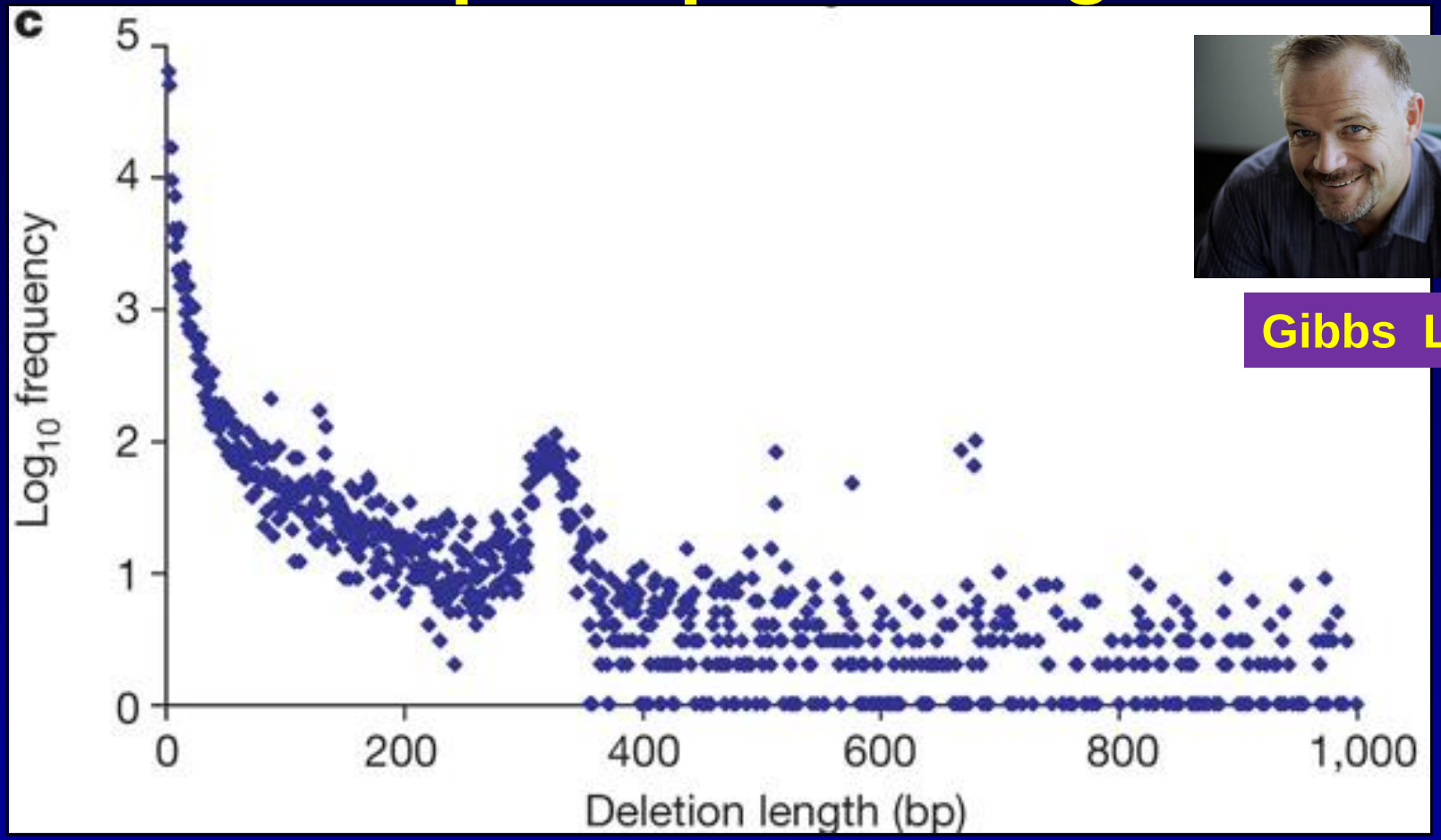
What did the Watson genome teach us?

Explaining clinical genetic implications of his personal genome to J.D. Watson

What did the Watson genome teach us?

1) First personal genome by NGS
2) Tremendous variation!
3) Millions of bases, no match to ref
4) Both SNV & CNV vary a lot!
5) Allele frequency spectrum of CNV reveals smaller more freq
6) Challenging to interpret

Size distribution of deletions in the Watson diploid personal genome.



Gibbs Lab

-Deletions observed in alignments of 454-reads to haploid human genome ref.
- Note peak at 300-350 bases owing to polymorphic Alu transposon dimorphisms. (i.e.insertion /deletion alleles) Wheeler, *et al.* (2008) *Nature 452*:872-876.

Personal Human Genomes Comparison

Individual	Ploidy	Technology	Av Depth	Total SNPs (M)	Known SNPs (M)	Novel SNPs (M)	Unique novel (M)
Venter	2n	Sanger	7.5x	3.21	2.80	0.74	0.52
Watson	2n	Roche 454	7.4x	3.32	2.71	0.61	0.57
Chinese (YH)	2n	Illumina/Solexa	36x	3.07	2.67	0.39	0.20
African (NA18507)*	2n	Illumina/Solexa	40.6x	3.61	2.72	0.88	0.52
African (NA18507)*	2n	AB SOLiD	17.9x	3.86	3.13	0.73	NA
Korean (SJK)	2n	Illumina/Solexa	28.95x	3.43	3.01	0.42	0.27
Korean (AK1)	2n	Illumina/Solexa	27.8x	3.45	2.88	0.57	NA
Neuropathy subject	2n	AB SOLiD	29.6x	3.42	2.85	0.56	0.46

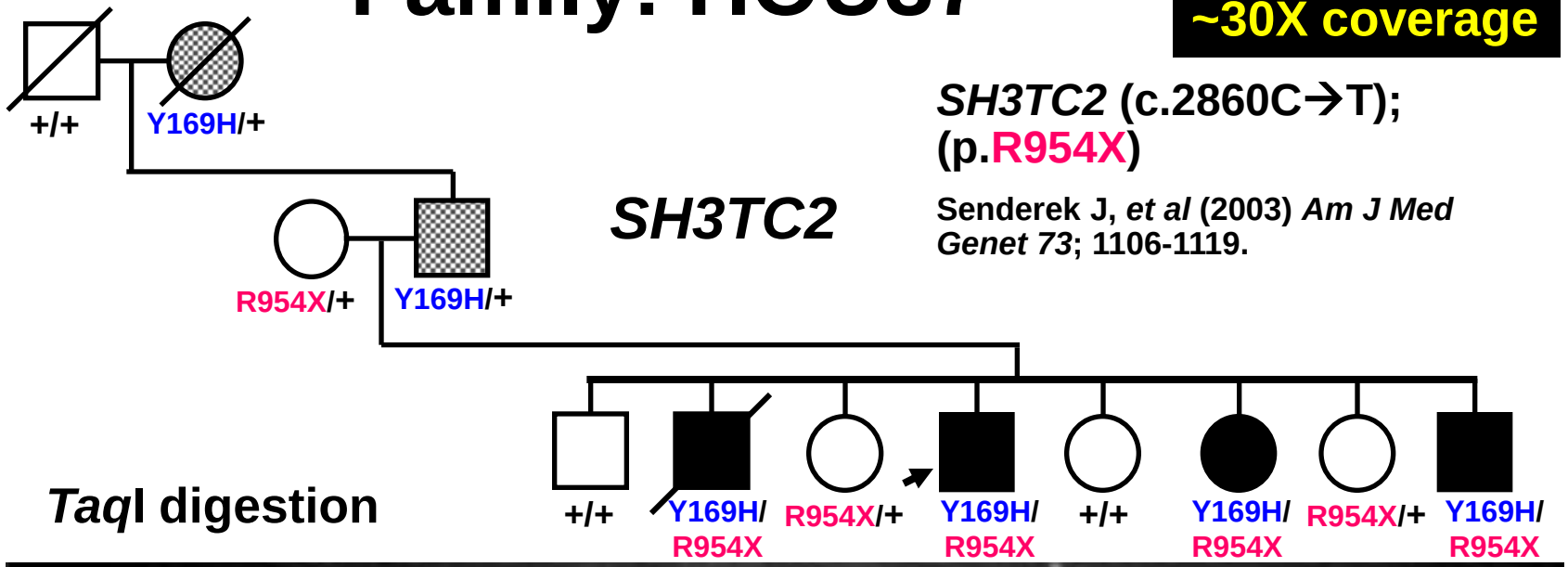
Associations of non-synonymous SNPs in Neuropathy subject's genome

	# of SNPs	Percentage
Total	*155*	*100%*
Behavioral Disorder	6	4%
Cancer Associated	32	21%
•Association	6	4%
•Increased risk	9	6%
•Reduced risk	3	2%
•Susceptibility	14	9%
Complex Disease	47	30%
Mendelian Disease	19	12%
Metabolic Trait	17	11%
Pharmacological Trait	14	9%
Other Traits	20	13%

Claudia Gonzaga-Jauregui

Family: HOU37

Proband @ ~30X coverage



SH3TC2

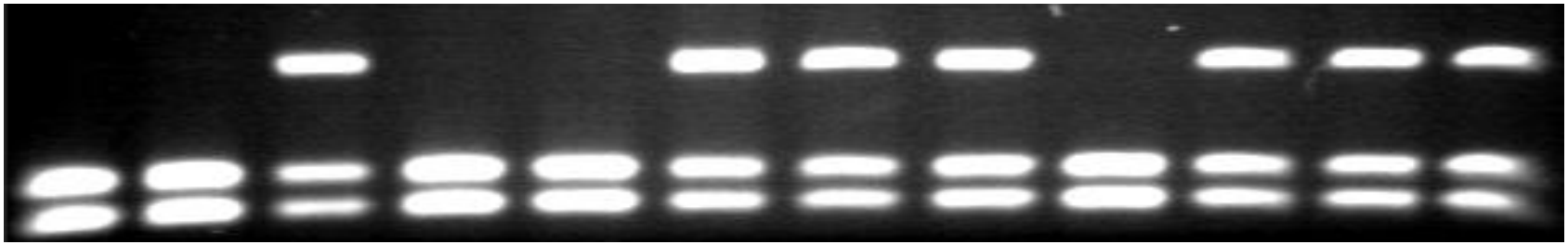
***SH3TC2* (c.2860C→T); (p.R954X)**

Senderek J, *et al* (2003) *Am J Med Genet 73*; 1106-1119.

***Taq*I digestion**

G→A
chr5: 148,386,628nt

WT

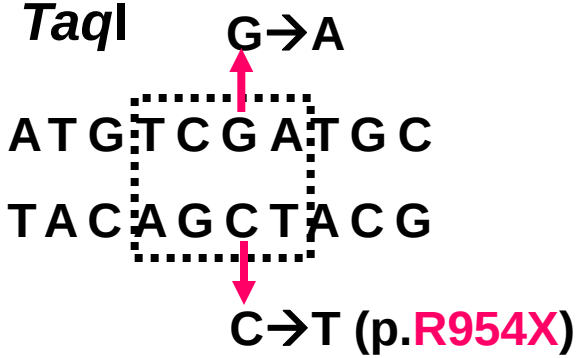


Y169

H sapiens	LVEDTEIQVSVDDKHLETIYLGLLIQEGHFFCR
P troglodytes	LVEDTEIQVSVDDKHLETIYLGLLIQEGHFFCR
M mulatta	LVEDTEIQVSVDDKHLETIYLGLLIQEGHFFCR
C familiaris	LVEDTEIQVSVDEKHLETIYLALLIQEGHFFCR
E caballus	LVEDTEIQVSVDDKHLESIYLGLLIQEGHFFCR
B taurus	LVEDTEIHVSIDDKHLETIYLGLLIQEGHFFCR
M musculus	LVDDTEIQVSVDDNHLENIYLGLLLQEGHFFCR
R norvegicus	LVDDTEIQVSVDDTHLENIYLGLLLQEGHFFCR
M domestica	LVEDTKIQVIVNYEHLEAIYQSLLIQEGH-FCR
G gallus	LVEDTEIRVSMDENRLATIYLGLLLQEGHFFSR

chr5:nt 148,402,474(A→G);

***Taq*I**

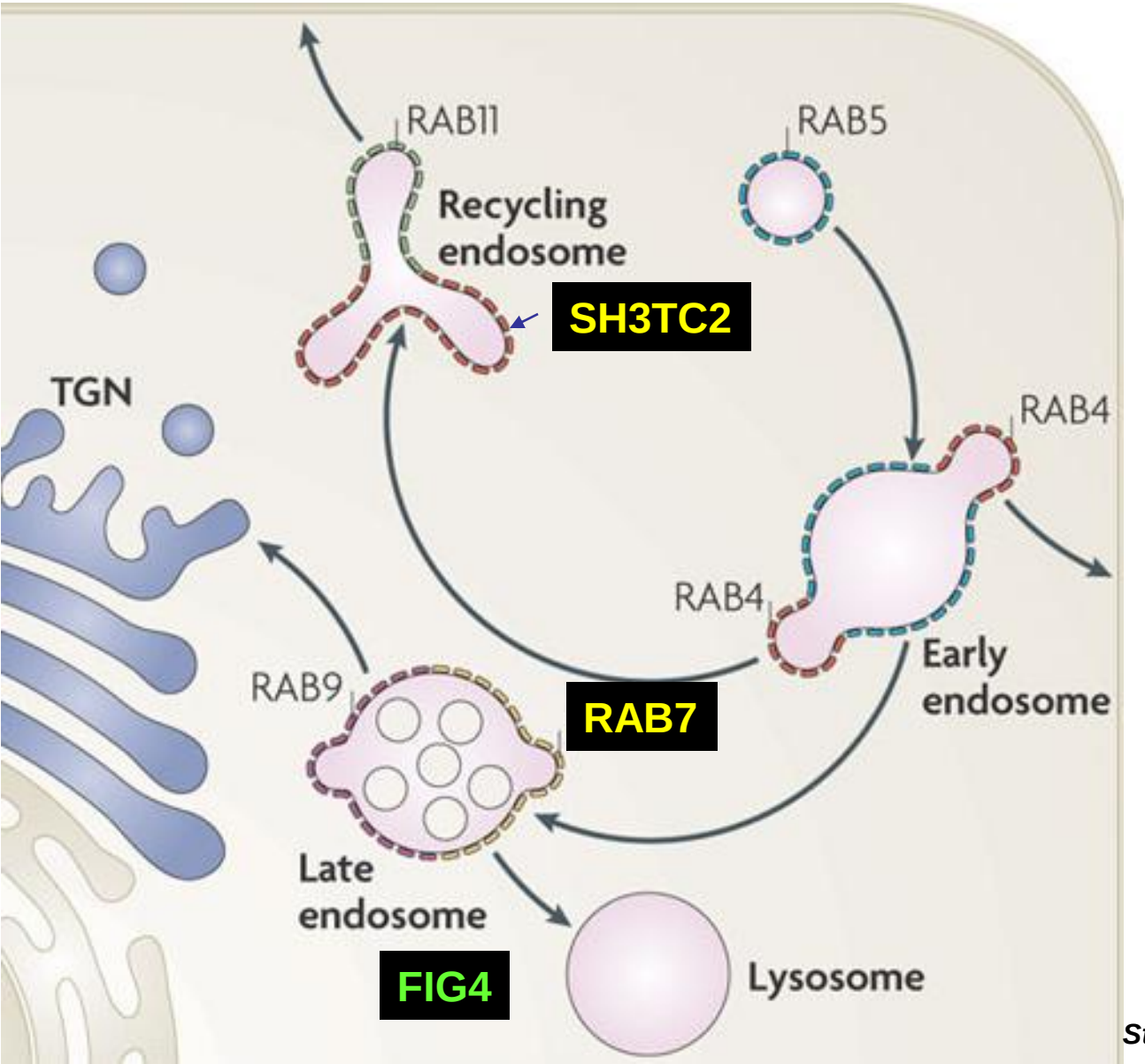


Claudia Gonzaga-Jauregui

SH3TC2 mutations cause AR CMT4C

- SH3TC2 protein contains SH3 and TPR motifs
- TPR mediate assembly of protein complexes binding to proline-rich proteins
- CMT4C associated patient mutations implicate endocytic and membrane trafficking pathway
- Sh3tc2 expressed in Schwann cells localizes to plasma membrane
- *Sh3tc2*$^{-/-}$ show abnormal node of Ranvier organization
- Possible function in myelination and/or in regions of axon-glia interactions
- Recent data suggest involvement in endocytic recycling

SH3TC2 plays a role in the endocytic recycling pathway



Pathogenic missense and nonsense mutations in *SH3TC2* apparently cause failure to localize to the recycling endosome and associate with Rab11.

Roberts, RC,* et. al, *Hum Mol Genet, 2010.

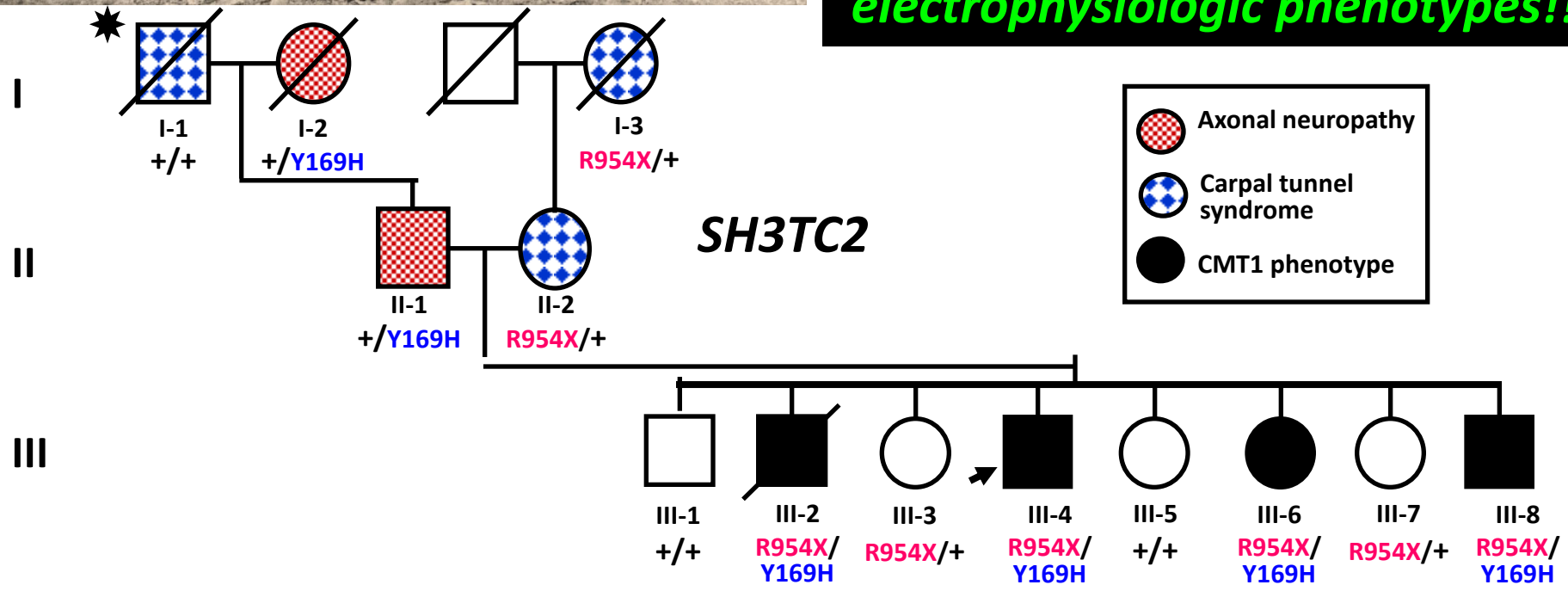
Stenmark, H., Nat Rev Mol Cell Biol, 2009.

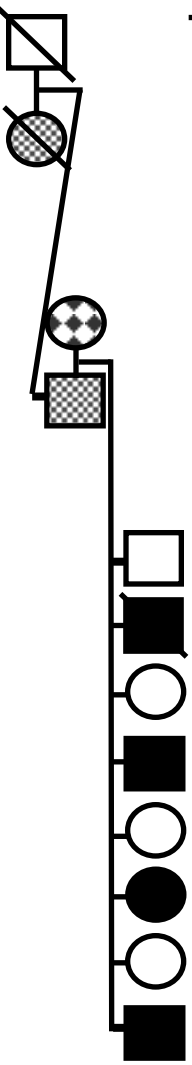
acquired

versus

inherited neuropathy

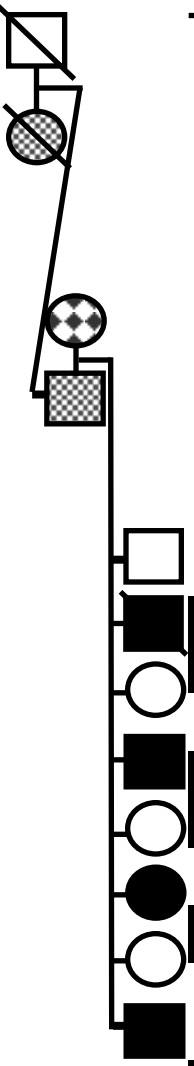
NCV studies distinguish three electrophysiologic phenotypes!!!





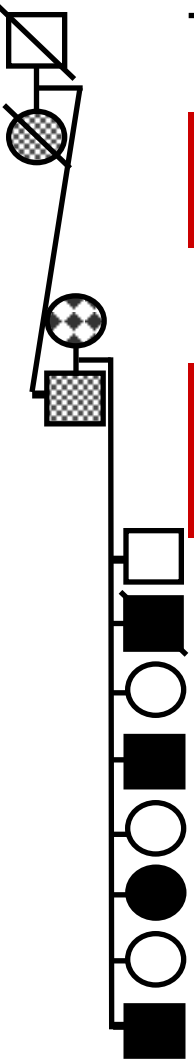
SUBJECT	AGE	SEX	DEMYELINATING	AXONAL	CTS MEDIAN NERVE	DX
I-1*	80	m	No	peroneal motor 2.2 mV.	y	
I-2	77	f	No	sural abs, motor: peroneal 0.5 mV., tibial 2.8 mV	y	axonal
I-3	90	f	No	otherwise nl	y SCV 43m/s	mmm
II-2	58	f	No	otherwise nl	y SCV 46m/s	mmm
II-1	57	m	No	sural abs, motor peroneal 0.2 mV, tibial 1.4 mV; H38ms	y	axonal
III-1	37	m	No	normal	n	
III-2	35	m	Yes	No	y term lat 14.9m/s median vs. 8.1 ulnar	CMT
III-3	34	f	No	No	y SCV 42 m/s; term lat 4.4	mmm
III-4	32	m	Yes	No	probably; term lat 10.2 median vs. 7.5 ulnar	CMT
III-5	31	f	No	No	n	
III-6	29	f	Yes	No	y term lat 11.6 median vs. 6.2 ulnar	CMT
III-7	26	f	No	peroneal 36 m/s; H reflexes 35ms	y SCV 36 m/s; term lat 4.8	mmm
III-8	25	m	Yes	No	y term lat 9.2 median vs. 6.2 ulnar	CMT

Abbreviations: mmm mild median mononeuropathy; CTS carpal tunnel syndrome; y yes; n no; abs: no response elicited; term lat; motor terminal latency (ms); SCV: sensory conduction velocity (m/s)
*was a carpenter for >50 years.



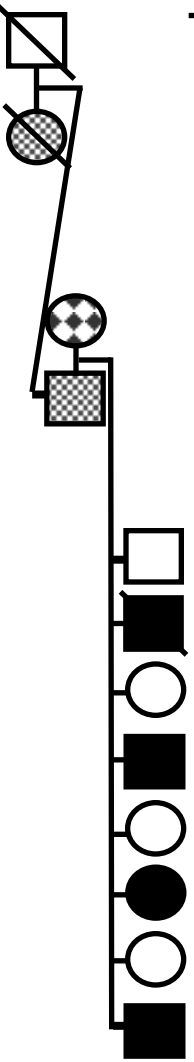
SUBJECT	AGE	SEX	DEMYELINATING	AXONAL	CTS MEDIAN NERVE	DX
I-1*	80	m	No	peroneal motor 2.2 mV.	y	
I-2	77	f	No	sural abs, motor: peroneal 0.5 mV., tibial 2.8 mV	y	axonal
I-3	90	f	No	otherwise nl	y SCV 43m/s	mmm
II-2	58	f	No	otherwise nl	y SCV 46m/s	mmm
II-1	57	m	No	sural abs, motor peroneal 0.2 mV, tibial 1.4 mV; H38ms	y	axonal
III-1	37	m	No	normal	n	
III-2	**35**	**m**	**Yes**	**No**	**y term lat 14.9m/s median vs. 8.1 ulnar**	**CMT**
III-3	34	f	No	No	y SCV 42 m/s; term lat 4.4	mmm
III-4	**32**	**m**	**Yes**	**No**	**probably; term lat 10.2 median vs. 7.5 ulnar**	**CMT**
III-5	31	f	No	No	n	
III-6	**29**	**f**	**Yes**	**No**	**y term lat 11.6 median vs. 6.2 ulnar**	**CMT**
III-7	26	f	No	peroneal 36 m/s; H reflexes 35ms	y SCV 36 m/s; term lat 4.8	mmm
III-8	**25**	**m**	**Yes**	**No**	**y term lat 9.2 median vs. 6.2 ulnar**	**CMT**

Abbreviations: mmm mild median mononeuropathy; CTS carpal tunnel syndrome; y yes; n no; abs: no response elicited; term lat; motor terminal latency (ms); SCV: sensory conduction velocity (m/s)
*was a carpenter for >50 years.



SUBJECT	AGE	SEX	DEMYELINATING	AXONAL	CTS MEDIAN NERVE	DX
I-1*	80	m	No	peroneal motor 2.2 mV.	y	
I-2	**77**	**f**	**No**	**sural abs, motor: peroneal 0.5 mV., tibial 2.8 mV**	**y**	**axonal**
I-3	90	f	No	otherwise nl	y SCV 43m/s	mmm
II-2	58	f	No	otherwise nl	y SCV 46m/s	mmm
II-1	**57**	**m**	**No**	**sural abs, motor peroneal 0.2 mV, tibial 1.4 mV; H38ms**	**y**	**axonal**
III-1	37	m	No	normal	n	
III-2	35	m	Yes	No	y term lat 14.9m/s median vs. 8.1 ulnar	CMT
III-3	34	f	No	No	y SCV 42 m/s; term lat 4.4	mmm
III-4	32	m	Yes	No	probably; term lat 10.2 median vs. 7.5 ulnar	CMT
III-5	31	f	No	No	n	
III-6	29	f	Yes	No	y term lat 11.6 median vs. 6.2 ulnar	CMT
III-7	26	f	No	peroneal 36 m/s; H reflexes 35ms	y SCV 36 m/s; term lat 4.8	mmm
III-8	25	m	Yes	No	y term lat 9.2 median vs. 6.2 ulnar	CMT

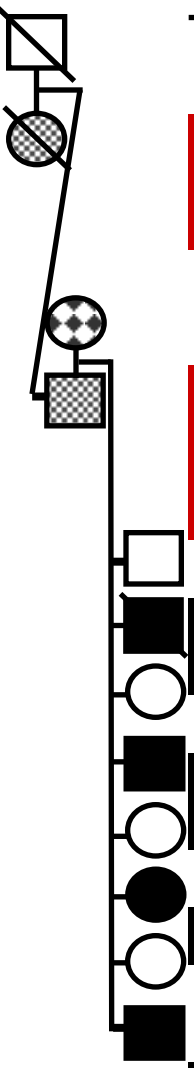
Abbreviations: mmm mild median mononeuropathy; CTS carpal tunnel syndrome; y yes; n no; abs: no response elicited; term lat; motor terminal latency (ms); SCV: sensory conduction velocity (m/s)
*was a carpenter for >50 years.



SUBJECT	AGE	SEX	DEMYELINATING	AXONAL	CTS MEDIAN NERVE	DX
I-1*	80	m	No	peroneal motor 2.2 mV.	**y**	
I-2	77	f	No	sural abs, motor: peroneal 0.5 mV., tibial 2.8 mV	**y**	axonal
I-3	90	f	No	otherwise nl	**y SCV 43m/s**	mmm
II-2	58	f	No	otherwise nl	**y SCV 46m/s**	mmm
II-1	57	m	No	sural abs, motor peroneal 0.2 mV, tibial 1.4 mV; H38ms	**y**	axonal
III-1	37	m	No	normal	n	
III-2	35	m	Yes	No	**y term lat 14.9m/s median vs. 8.1 ulnar**	CMT
III-3	34	f	No	No	**y SCV 42 m/s; term lat 4.4**	
III-4	32	m	Yes	No	**probably; term lat 10.2 median vs. 7.5 ulnar**	CMT
III-5	31	f	No	No	n	
III-6	29	f	Yes	No	**y term lat 11.6 median vs. 6.2 ulnar**	CMT
III-7	26	f	No	peroneal 36 m/s; H reflexes 35ms	**y SCV 36 m/s; term lat 4.8**	
III-8	25	m	Yes	No	**y term lat 9.2 median vs. 6.2 ulnar**	CMT

Abbreviations: mmm mild median mononeuropathy; CTS carpal tunnel syndrome; y yes; n no; abs: no response elicited; term lat; motor terminal latency (ms); SCV: sensory conduction velocity (m/s)

*was a carpenter for >50 years



SUBJECT	AGE	SEX	DEMYELINATING	AXONAL	CTS MEDIAN NERVE	DX
I-1*	80	m	No	peroneal motor 2.2 mV.	y	
I-2	**77**	**f**	**No**	**sural abs, motor: peroneal 0.5 mV., tibial 2.8 mV**	**y**	**axonal**
I-3	90	f	No	otherwise nl	y SCV 43m/s	mmm
II-2	58	f	No	otherwise nl	y SCV 46m/s	mmm
II-1	**57**	**m**	**No**	**sural abs, motor peroneal 0.2 mV, tibial 1.4 mV; H38ms**	**y**	**axonal**
III-1	37	m	No	normal	n	
III-2	**35**	**m**	**Yes**	**No**	**y term lat 14.9m/s median vs. 8.1 ulnar**	**CMT**
III-3	34	f	No	No	y SCV 42 m/s; term lat 4.4	mmm
III-4	**32**	**m**	**Yes**	**No**	**probably; term lat 10.2 median vs. 7.5 ulnar**	**CMT**
III-5	31	f	No	No	n	
III-6	**29**	**f**	**Yes**	**No**	**y term lat 11.6 median vs. 6.2 ulnar**	**CMT**
III-7	26	f	No	peroneal 36 m/s; H reflexes 35ms	y SCV 36 m/s; term lat 4.8	mmm
III-8	**25**	**m**	**Yes**	**No**	**y term lat 9.2 median vs. 6.2 ulnar**	**CMT**

Abbreviations: mmm mild median mononeuropathy; CTS carpal tunnel syndrome; y yes; n no; abs: no response elicited; term lat; motor terminal latency (ms); SCV: sensory conduction velocity (m/s)
*was a carpenter for >50 years.

The NEW ENGLAND JOURNAL of MEDICINE

ORIGINAL ARTICLE

Whole-Genome Sequencing in a Patient with Charcot–Marie–Tooth Neuropathy

James R. Lupski, M.D., Ph.D., Jeffrey G. Reid, Ph.D., Claudia Gonzaga-Jauregui, B.S., David Rio Deiros, B.S., David C.Y. Chen, M.Sc., Lynne Nazareth, Ph.D., Matthew Bainbridge, M.Sc., Huyen Dinh, B.S., Chyn Jing, M.Sc., David A. Wheeler, Ph.D., Amy L. McGuire, J.D., Ph.D., Feng Zhang, Ph.D., Pawel Stankiewicz, M.D., Ph.D., John J. Halperin, M.D., Chengyong Yang, Ph.D., Curtis Gehman, Ph.D., Danwei Guo, M.Sc., Rola K. Irikat, B.S., Warren Tom, B.S., Nick J. Fantin, B.S., Donna M. Muzny, M.Sc., and Richard A. Gibbs, Ph.D.

***362* :1181-1191 (2010)**

ABSTRACT

BACKGROUND

Whole-genome sequencing may revolutionize medical diagnostics through rapid identification of alleles that cause disease. However, even in cases with simple patterns of inheritance and unambiguous diagnoses, the relationship between disease phenotypes and their corresponding genetic changes can be complicated. Comprehensive diagnostic assays must therefore identify all possible DNA changes in each haplotype and determine which are responsible for the underlying disorder. The high number of rare, heterogeneous mutations present in all humans and the paucity of known functional variants in more than 90% of annotated genes make this challenge particularly difficult. Thus, the identification of the molecular basis of a genetic disease by means of whole-genome sequencing has remained elusive. We therefore aimed to assess the usefulness of human whole-genome sequencing for genetic diagnosis in a patient with Charcot–Marie–Tooth disease.

METHODS

We identified a family with a recessive form of Charcot–Marie–Tooth disease for which the genetic basis had not been identified. We sequenced the whole genome of the proband, identified all potential functional variants in genes likely to be related to the disease, and genotyped these variants in the affected family members.

RESULTS

We identified and validated compound, heterozygous, causative alleles in *SH3TC2* (the SH3 domain and tetratricopeptide repeats 2 gene), involving two mutations, in the proband and in family members affected by Charcot–Marie–Tooth disease. Separate subclinical phenotypes segregated independently with each of the two mutations; heterozygous mutations confer susceptibility to neuropathy, including the carpal tunnel syndrome.

CONCLUSIONS

As shown in this study of a family with Charcot–Marie–Tooth disease, whole-genome sequencing can identify clinically relevant variants and provide diagnostic information to inform the care of patients.

"All the News That's Fit to Print"

The New York Times

National Edition

Gulf Coast: Partly cloudy. Rain early in the northeast. Highs in 50s in the northwest, 70s in the south. Showers tonight, in the north and east. Weather map is on Page B13.

VOL. CLIX . . . . No. 54,976 © 2010 The New York Times THURSDAY, MARCH 11, 2010 Printed in Houston $2.00

Disease Cause Is Pinpointed With Genome

Decoding Cost Drops to $50,000 a Patient

By NICHOLAS WADE

Two research teams have independently decoded the entire genome of patients to find the exact genetic cause of their diseases. The approach may offer a new start in the so far disappointing effort to identify the genetic roots of major killers like heart disease, diabetes and Alzheimer's.

In the decade since the first full genetic code of a human was sequenced for some $500 million, less than a dozen genomes had been decoded, all of healthy people.

Geneticists said the new research showed it was now possible to sequence the entire genome of a patient at reasonable cost and with sufficient accuracy to be of practical use to medical researchers. One subject's genome cost just $50,000 to decode.

"We are finally about to turn the corner, and I suspect that in the next few years human genetics will finally begin to systematically deliver clinically meaningful findings," said David B. Goldstein, a Duke University geneticist who has criticized the current approach to identifying genetic causes of common diseases.

Whole-Genome Sequencing Offers Clues to Diseases

From Page A1

sequenced the whole genome of his colleague Dr. James R. Lupski, a prominent medical geneticist who has a nerve disease, Charcot-Marie-Tooth neuropathy.

In the second, Leroy Hood and David J. Galas of the Institute for Systems Biology in Seattle have decoded the genomes of two children with two rare genetic diseases, and their parents.

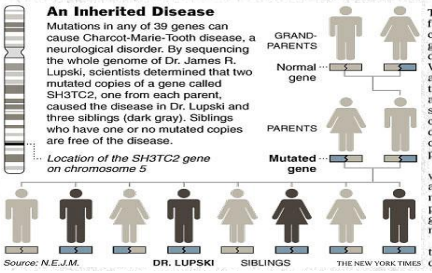
More common diseases, like cancer, are thought to be caused by mutations in several genes, and finding the causes was the principal goal of the $3 billion human genome project. To that end, medical geneticists have invested heavily over the last eight years in an alluring shortcut.

But the shortcut was based on a premise that is turning out to be incorrect. Scientists thought the mutations that caused common diseases would themselves be common. So they first identified the common mutations in the human population in a $100 million project called the HapMap. Then they compared patients' genomes with those of healthy genomes. The comparisons relied on ingenious devices called SNP chips, which scan just a tiny portion of the genome. (SNP, pronounced "snip," stands for single nucleotide polymorphism.) These projects, called genome-wide association studies, each cost around $10 million or more.

The results of this costly international exercise have been disappointing. About 2,000 sites on the human genome have been statistically linked with various diseases, but in many cases the sites are not inside working genes, suggesting there may be some conceptual flaw in the statistics. And in most diseases the

MICHAEL STRAVATO FOR THE NEW YORK TIMES

Dr. James R. Lupski, a medical geneticist with a nerve disease, had his whole genome decoded.



They analyzed the genome of a family of four, in which the two children each have two single-gene diseases, called Miller syndrome and ciliary dyskinesia. With four related genomes available, the researchers could identify the causative genes. They also improved the accuracy of the sequencing, because DNA changes that did not obey Mendel's rules of inheritance could be classed as errors in the decoding process.

The Seattle team believes whole-genome sequencing can be applied to the study of the common multigene diseases and plans to sequence more than 100 genomes next year, starting with multigenerational families.

The family whose genomes they report in Science were sequenced by a company with a

BCM
Baylor College of Medicine

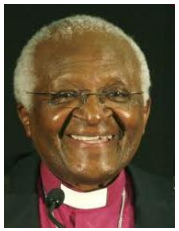
HGSC
HUMAN GENOME SEQUENCING CENTER

Richard A. Gibbs

Whole-genome sequencing (WGS):
milestones in the path to personal medical genomics

James D. Watson
Nature (2008)452: 872-876.

→ **Individual variation**



Desmond Tutu
(Khoisan & Bantu genomes)
Nature (2010) 463:943-947.

→ **Population variation**

Jim Lupski
New Engl J Med (2010)362:1181-1191.

→ **Identification of medically actionable, disease-causing variants**

Beery twins
Sci Transl Med (2011) 3(87);87re3.

→ **Medical management & therapeutic modifications based on personal variation**

CONCLUSIONS-What have we learned?

- Rare variants, genetic heterogeneity, total mutational load (SNV + CNV; inherited + *de novo*) can explain at least some common and complex traits
- To what extent will 'exon dropout CNV' and 'CNV of non-coding regions' account for "missing heritability"
- WGS can identify causative alleles for a Mendelian AR trait; CMT4F
- WGS can inform clinical observations - mis vrnt segregates with NCV observed axonal neuropathy in CMT family.
- *SH3TC2 & PMP22* haploinsufficiency, the latter via HNPP del, can confer susceptibility to the carpal tunnel syndrome - a complex trait
- WGS may be a cost effective way to screen for mutant alleles in a very genetically heterogeneous trait (e.g. deafness, retinitis pigmentosa, mental retardation, CMT, etc.)

Contributors

Personal genome sequencing


Matthew Bainbridge


David Murdock


Jeffrey Reid


Richard Gibbs


Claudia Gonzaga-Jauregui


Feng Zhang


Pawel Stankiewicz


Jim Lupski

+ Several Investigators from: Life Technologies, Inc

ACKNOWLEDGEMENTS:

Gibbs Lab & Baylor

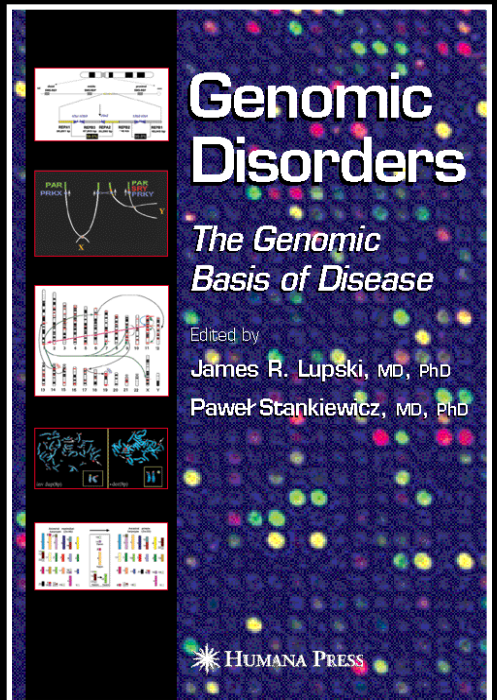
HGSC
HUMAN GENOME SEQUENCING CENTER

+ Lupski Lab

&

MGL
MEDICAL GENETICS LABORATORIES

WGL
WHOLE GENOME LABORATORY



E-mail: jlupski@bcm.edu

http://imgen.bcm.tmc.edu/molgen/lupski/