Ogden Syndrome and the Amino-Terminal Acetylation of Proteins

Gholson J. Lyon, M.D. Ph.D.

For

NIGMS Medical Scientist Training Program 50th Anniversary Symposium

Cold Spring Harbor Laboratory
Stanley Institute for Cognitive Genomics

Acknowledgments

Martin Reese
Edward Kiruluta

Barry Moore
Alan Rope
Jeffrey J Swensen
Lynn Jorde
Mark Yandell

Thomas Arnesen
Nathalie Reuter
Line Myklebust

Reid Robison

Jason O'Rawe
Yiyang Wu
Han Fang
Max Doerfel
Michael Schatz
Giuseppe Narzisi

our study families
and many others

David Mittelman

Kai Wang

Tina Hambuch
Erica Davis
Dawn Barry

Ghent, Belgium
Petra Van Damme
Kris Gevaert

Jason O’Rawe

Yiyang Wu

Han Fang

Max Doerfel

Involvement with Industry

Advisory Boards

genepeeks

micia
UNLOCKING INDIVIDUALIZED MEDICINE

Other non-paid consulting:

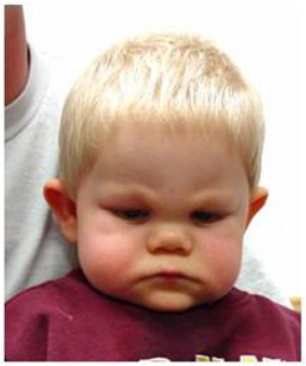
GOLDEN HELIX
Accelerating the Quest for Significance™

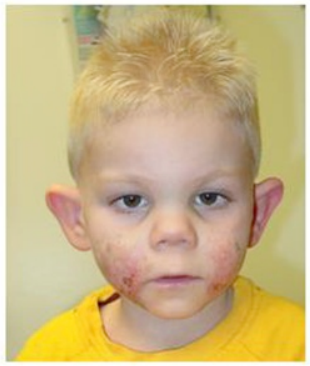
TUTE GENOMICS

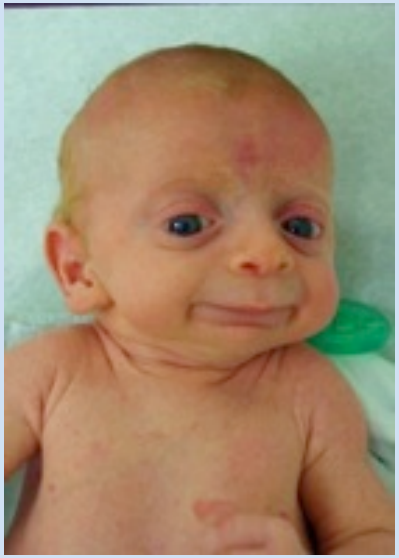
The Big Picture

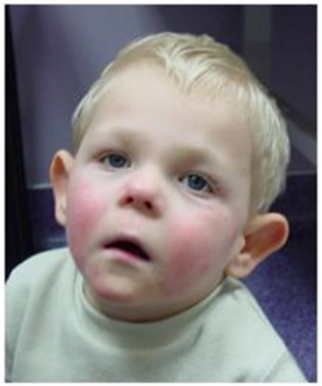
- Over the course of my entire career, I want to help understand the pathophysiology of severe disorders with neuropsychiatric manifestations, including such things as developmental delay, intellectual disability, and other neurobehavioral outcomes.
- This will certainly uncover new biology along the way.

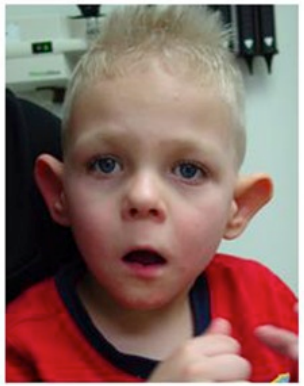
Rare Diseases

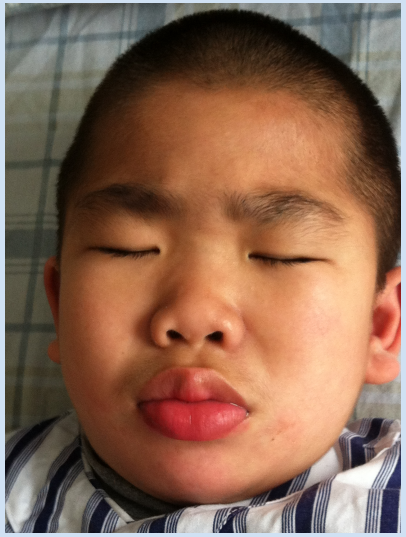


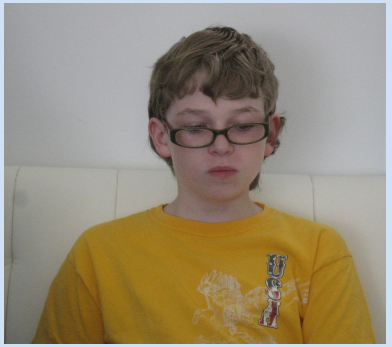






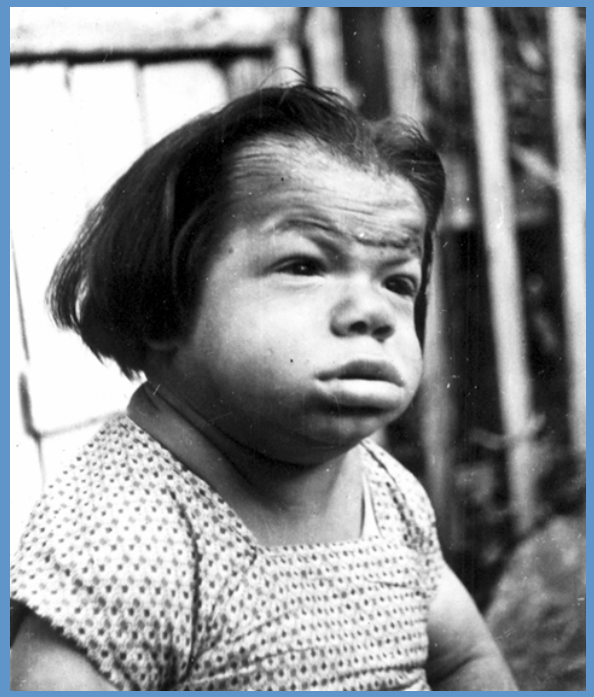


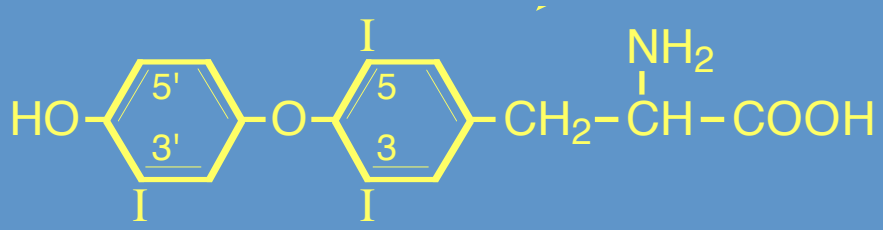




The Story began for me at least by 1993….

when I studied, as an undergraduate student, the role of thyroid hormone in cretinism, which is caused by lack of iodine during maternal pregnancy, so this is an environmentally triggered disease, modified by genetic and other factors.





Thyroid Hormone

1996-97

- In Cambridge, England at Christ's College, worked with Martin Evans on mouse knockouts and models of human disease.
- Met Alexander Bearn, a distinguished human geneticist who happened to be a fellow in residence at the time. He wrote the definitive biography of Archibald Garrod.
- At that time, I formalized my goal to study human genetics long-term.

What will it take for me to study human genetics and certain diseases in detail over my entire career?

"The M.D. does not make you a physician,
but it prepares you to become one.

The Ph.D. does not make you a scientist,
but it prepares you to become one."

-Olaf Andersen, M.D.
Director of Cornell/Rockefeller/Sloan-Kettering
M.D./Ph.D. program

M.D. Ph.D. training 1997-2004

- Much learning of human physiology, anatomy and disease at Weill Cornell Medical College.

- Ph.D. in bacterial genetics and chemical biology with Tom Muir at Rockefeller and Richard Novick at NYU.

2004- present

Becoming a clinician through clinical residency and practice.

Becoming a scientist through focusing on the pathophysiology of idiopathic syndromes.

2009- present

THE UNIVERSITY OF UTAH

UFBR
UTAH FOUNDATION FOR BIOMEDICAL RESEARCH

Discovering a new syndrome and its genetic basis.

ARTICLE

Using VAAST to Identify an X-Linked Disorder Resulting in Lethality in Male Infants Due to N-Terminal Acetyltransferase Deficiency

Alan F. Rope,[1] Kai Wang,[2,19] Rune Evjenth,[3] Jinchuan Xing,[4] Jennifer J. Johnston,[5] Jeffrey J. Swensen,[6,7] W. Evan Johnson,[8] Barry Moore,[4] Chad D. Huff,[4] Lynne M. Bird,[9] John C. Carey,[1] John M. Opitz,[1,4,6,10,11] Cathy A. Stevens,[12] Tao Jiang,[13,14] Christa Schank,[8] Heidi Deborah Fain,[15] Reid Robison,[15] Brian Dalley,[16] Steven Chin,[6] Sarah T. South,[1,7] Theodore J. Pysher,[6] Lynn B. Jorde,[4] Hakon Hakonarson,[2] Johan R. Lillehaug,[3] Leslie G. Biesecker,[5] Mark Yandell,[4] Thomas Arnesen,[3,17] and Gholson J. Lyon[15,18,20,*]

The American Journal of Human Genetics *89*, 1–16, July 15, 2011

Published online 23 June 2011 | Nature | doi:10.1038/news.2011.382

News

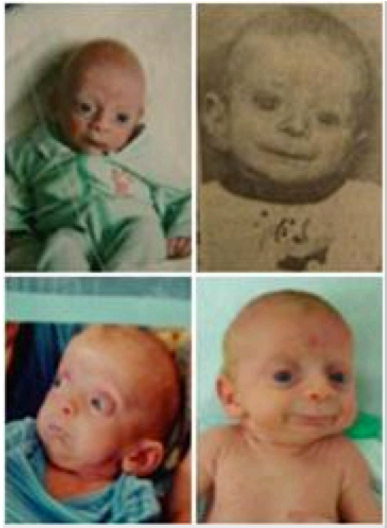
Software pinpoints cause of mystery genetic disorder

Genome analysis tools speedily track down previously unknown mutation.

Brendan Maher

Halena Black's first son, Kenny Rae, was born in November 1979. He struggled to put on weight, and had thin, wrinkled skin, big eyes and a broad mouth. In October the following year, he died from heart problems.

After Kenny Rae died, Black — a Mormon living in Ogden, Utah — had three healthy daughters before giving birth to another son in 1987. He had the same problem, and a similarly short lifespan. Her third son is healthy.



Four boys from the same family born wirh 'Ogden Syndrome'. Sufferers rarely survive for more than a year.

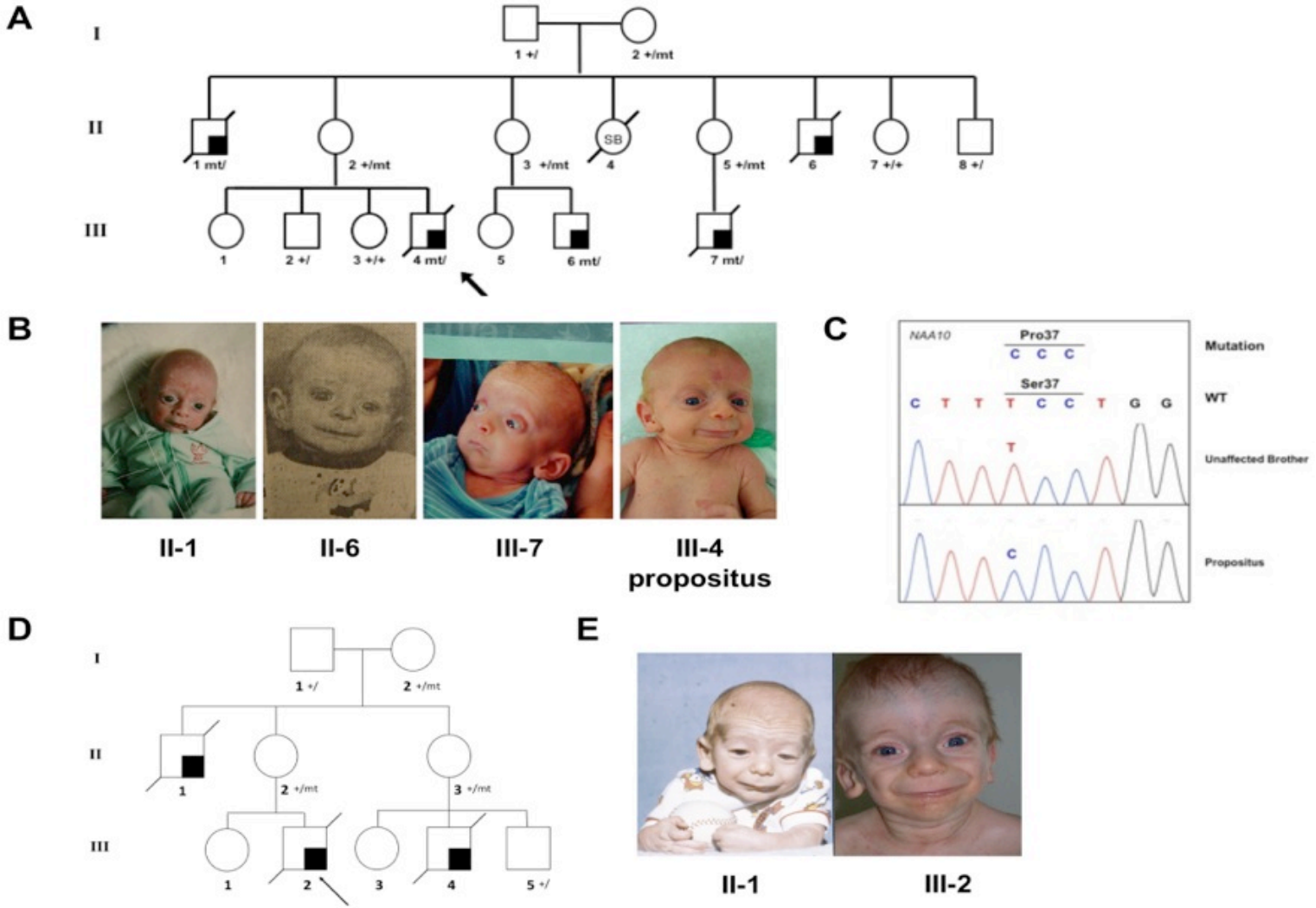
A. Rope et al./Am. J. Hum. Genet.

Black says she didn't dwell much on why her sons died until one of her daughters gave birth to a boy who looked just like Kenny Rae. "We didn't think that it passed on to the next generation. We didn't think that this would be a problem for them," says Black. All three daughters have since given birth to what the family calls 'little old men', one of whom died just last Sunday.

"This exemplifies an exceptionally rare disease, but the same type of strategy is now going to be applied to more common diseases to get the root cause," says Eric Topol, a medical geneticist at the Scripps Research Institute in La Jolla, California.

"This is one of the most exciting things in medicine," says Topol. "We're going to take the term 'idiopathic' which, basically means 'we don't' know,' and eliminate it."

Ogden Syndrome

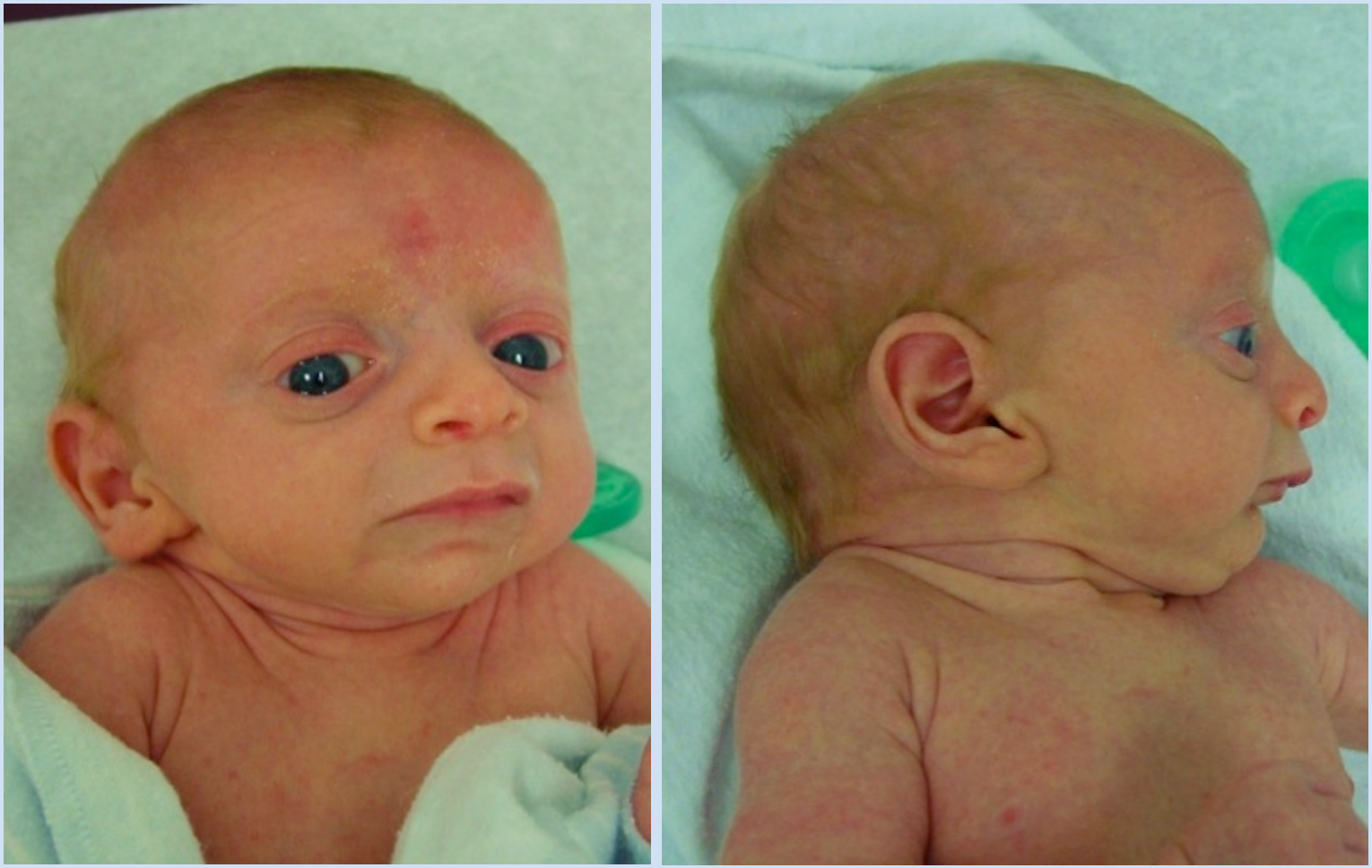


We found the SAME mutation in two unrelated families, with a very similar phenotype in both families, helping prove that this genotype contributes to the phenotype observed.

This is the first boy in the late 1970's.

First boy. Called "a little old man" by the family. Died around ~1 year of age, from cardiac arrhythmias.

This is the "Proband" photograph presented at Case Conference.



prominence of eyes, down-sloping palpebral fissures, thickened eyelids, large ears, beaking of nose, flared nares, hypoplastic nasal alae, short columella, protruding upper lip, micro-retrognathia

These are the Major Features of the Syndrome.

Table 1. Features of the syndrome	
Growth	post-natal growth failure
Development	global, severe delays
Facial	prominence of eyes, down-sloping palpebral fissures, thickened lids large ears beaking of nose, flared nares, hypoplastic alae, short columella protruding upper lip micro-retrognathia
Skeletal	delayed closure of fontanels broad great toes
Integument	redundancy / laxity of skin minimal subcutaneous fat cutaneous capillary malformations
Cardiac	structural anomalies (ventricular septal defect, atrial level defect, pulmonary artery stenoses) arrhythmias (Torsade de points, PVCs, PACs, SVtach, Vtach) death usually associated with cardiogenic shock preceded by arrythmia.
Genital	inguinal hernia hypo- or cryptorchidism
Neurologic	hypotonia progressing to hypertonia cerebral atrophy neurogenic scoliosis
Shaded regions include features of the syndrome demonstrating variability. Though variable findings of the cardiac, genital and neurologic systems were observed, all affected individuals manifested some pathologic finding of each.	

- **We performed X-chromosome exon capture with Agilent, followed by Next Gen Sequencing with Illumina.**

- **We analyzed the data with ANNOVAR and VAAST (Variant Annotation, Analysis and Search Tool). New computational tools for identifying disease-causing mutations by individual genome sequencing.**

Yandell, M. *et al.* 2011. "A probabilistic disease-gene finder for personal genomes." *Genome Res*. 21 (2011). doi:10.1101/gr.123158.111.

Wang, K., Li, M., and Hakonarson, H. (2010). ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res 38, e164.

VAAST integrates AAS & Variant frequencies in a single probabilistic framework

- non-coding variants scored using allele frequency differences

- n_i : frequency of variant type among all variants observed in Background and Target genomes

- a_i: frequency of variant type among disease causing mutations in OMIM

- This approach means that *every* variant can be scored, non-synonymous, synonymous, coding, and non-coding. Phylogenetic conservation not required.

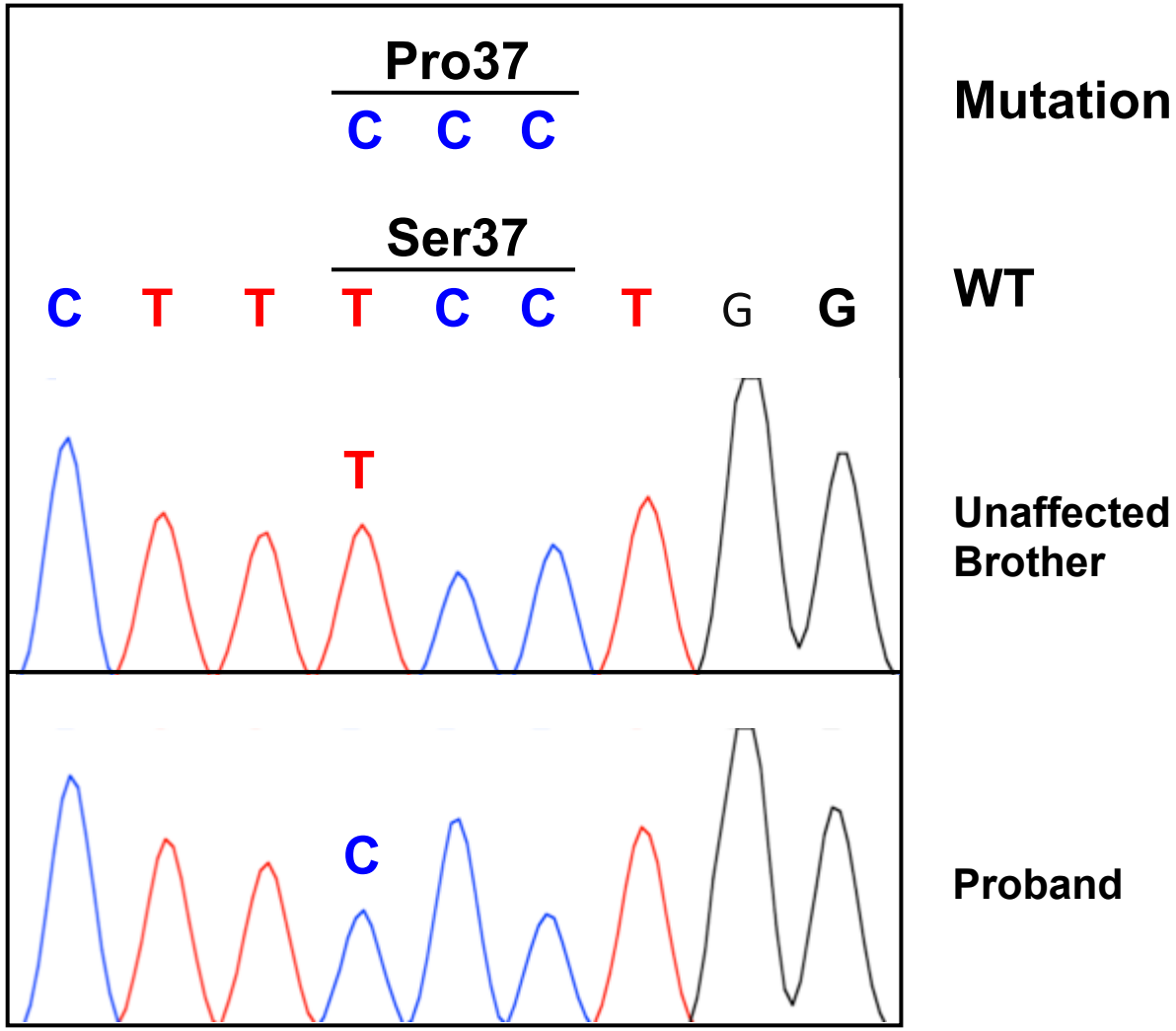
Omicia™ Unlocking Individualized Medicine

THE UNIVERSITY OF UTAH

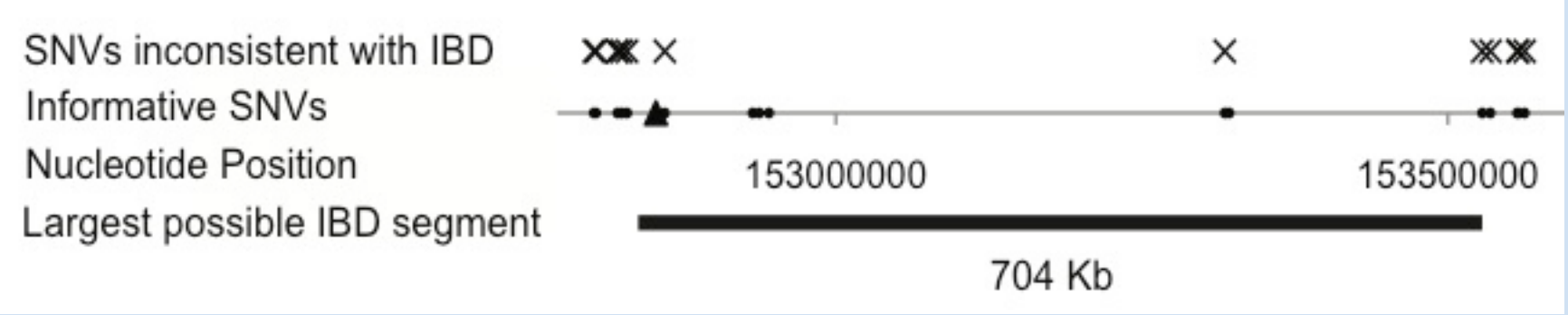
Analysis with VAAST readily identified a few likely candidates.

Table 3. Summary of the filtering procedure and candidate genes using VAAST			
SNV calling pipeline	GATK	Samtools	GNUMAP
III-4 (total SNVs)	1546	1499	2168
III-4 (nsSNVs)	146	114	155
VAAST candidate genes (NAA10 ranking)	4 (3)	3 (2)	5 (2)
Present in III-4 and mother II-2 (nsSNVs)	122	107	116
VAAST candidate genes (NAA10 ranking)	3 (2)	2 (1)	2 (2)
Present in III-4, mother II-2, and grandmother I-2 (nsSNVs)	115	95	104
VAAST candidate genes (NAA10 ranking)	2 (1)	2 (1)	1 (1)
Present in III-4, II-2, and I-2, absent in brother III-2 and uncle II-8 (nsSNVs)	8	6	8
VAAST candidate genes (NAA10 ranking)	1 (1)	1 (1)	2 (1)

This is the mutation we found… one nucleotide change out of 6 billion nucleotides in a diploid genome…

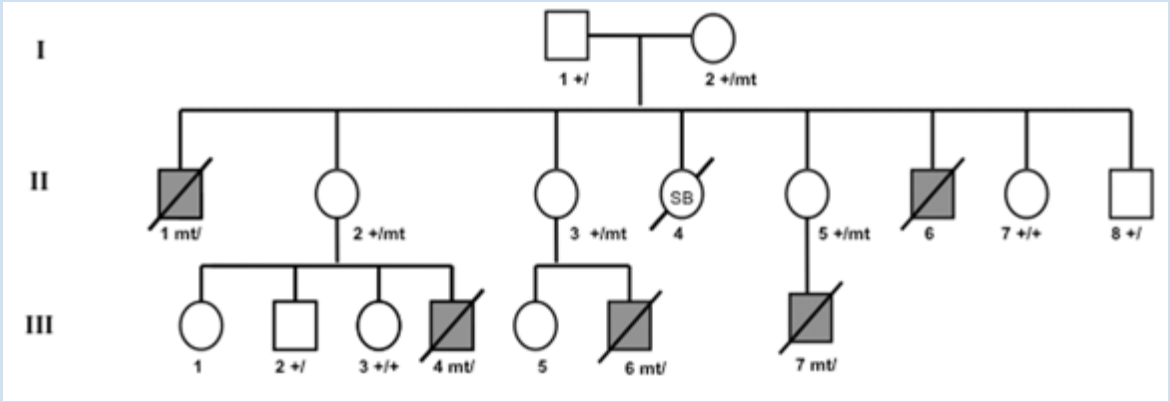


Identity by Descent Analysis shows that the mutation must have arisen independently in two different families.

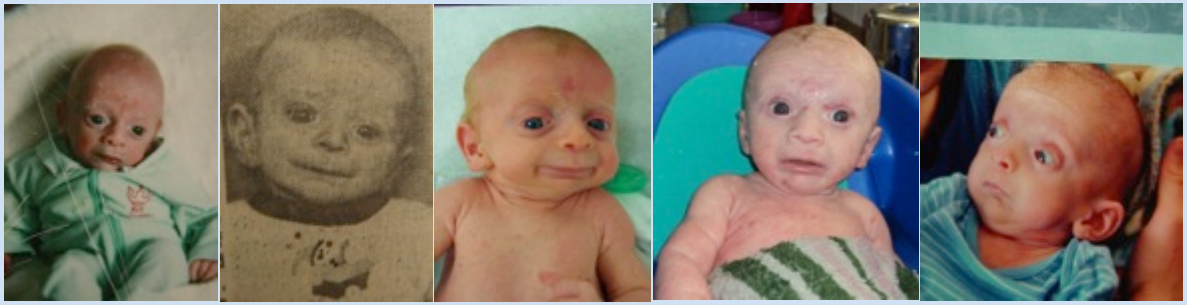


Courtesy of Chad Huff and Lynn Jorde

A

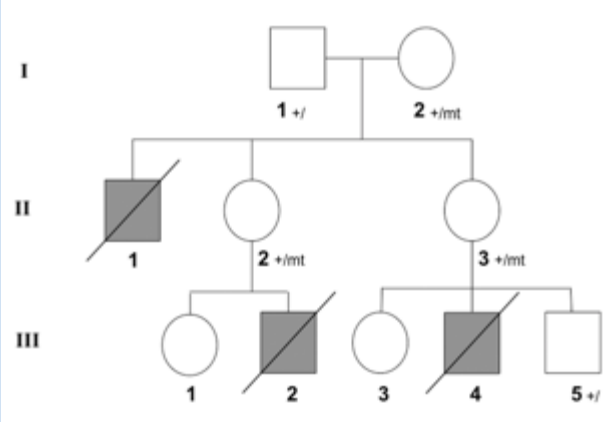


B

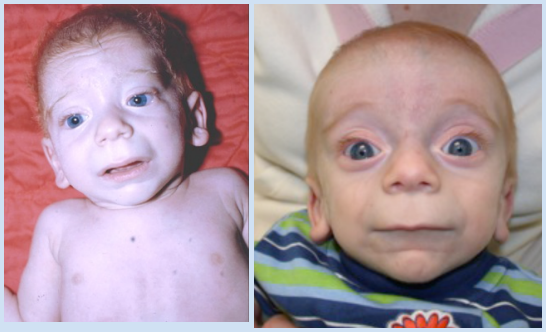


II-1 II-6 III-4 III-6 III-7

C



D



II-1 III-2

Moved to CSHL in 2012

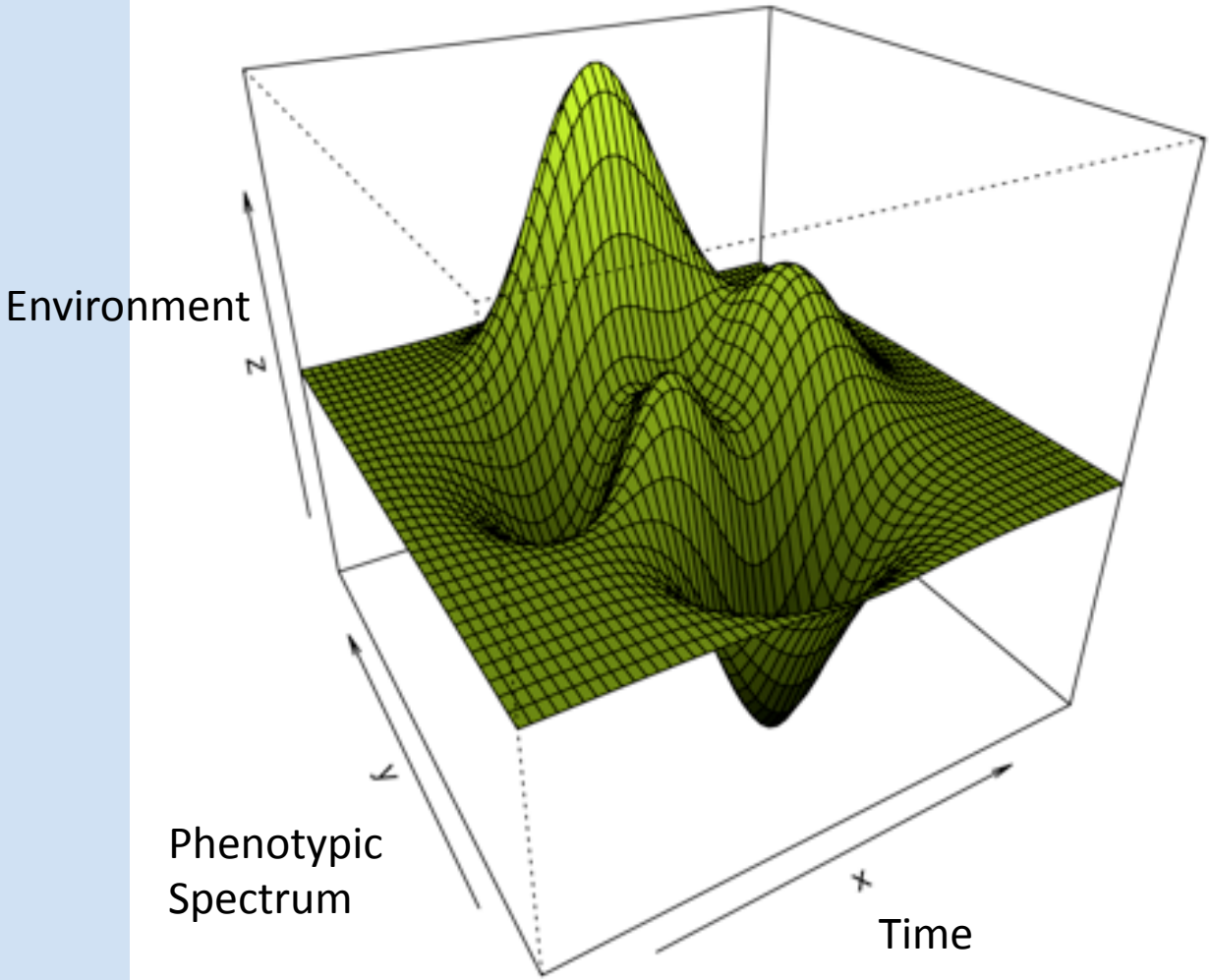
Clinical genetics of neurodevelopmental disorders

Gholson J Lyon and Jason O'Rawe

bioRxiv posted online November 18, 2013
Access the most recent version at doi:10.1101/000687

"There are ~12 billion nucleotides in every cell of the human body, and there are ~25-100 trillion cells in each human body. Given somatic mosaicism, epigenetic changes and environmental differences, no two human beings are the same, particularly as there are only ~7 billion people on the planet".

To appear as a book chapter: Lyon, G. J.* and O'Rawe J.A. Human genetics and clinical aspects of neurodevelopmental disorders. In: The Genetics of Neurodevelopmental Disorders, publisher: Wiley, 2014.



A conceptual model of genotype-phenotype correlations. The *y* plane represents a phenotypic spectrum, the *x* plane represents the canalized progression of development through time, and the *z* plane represents environmental fluctuations.

Prioritization of neurodevelopmental disease genes by discovery of new mutations

Alexander Hoischen[1], Niklas Krumm[2] & Evan E Eichler[2,3]

Advances in genome sequencing technologies have begun to revolutionize neurogenetics, allowing the full spectrum of genetic variation to be better understood in relation to disease. Exome sequencing of hundreds to thousands of samples from patients with autism spectrum disorder, intellectual disability, epilepsy and schizophrenia provides strong evidence of the importance of *de novo* and gene-disruptive events. There are now several hundred new candidate genes and targeted resequencing technologies that allow screening of dozens of genes in tens of thousands of individuals with high specificity and sensitivity. The decision of which genes to pursue depends on many factors, including recurrence, previous evidence of overlap with pathogenic copy number variants, the position of the mutation in the protein, the mutational burden among healthy individuals and membership of the candidate gene in disease-implicated protein networks. We discuss these emerging criteria for gene prioritization and the potential impact on the field of neuroscience.

Table 4 Recurrent identical *de novo* mutations in 6 genes identified in 11 exome studies with different neurodevelopmental phenotypes

Gene	Coding effect	Mutation (genomic DNA level)	Mutation (cDNA level)	Mutation (protein level)	Study	Disorder
ALG13	Missense	ChrX(GRCh37):g.110928268A>G	NM_001099922.2:c.320A>G	p.Asn107Ser	de Ligt *et al.*[1]	ID
ALG13	Missense	ChrX(GRCh37):g.110928268A>G	NM_001099922.2:c.320A>G	p.Asn107Ser	Allen *et al.*[11]	EE
ALG13	Missense	ChrX(GRCh37):g.110928268A>G	NM_001099922.2:c.320A>G	p.Asn107Ser	Allen *et al.*[11]	EE
KCNQ3	Missense	Chr8(GRCh37):g.133192493G>A	NM_001204824.1:c.328C>T	p.Arg110Cys	Rauch *et al.*[2]	ID
KCNQ3	Missense	Chr8(GRCh37):g.133192493G>A	NM_001204824.1:c.328C>T	p.Arg110Cys	Allen *et al.*[11]	EE
SCN1A	Splice donor	LRG_8:g.24003G>A	NM_006920.4:c.602+1G>A	p.?	Allen *et al.*[11]	EE
SCN1A	Splice donor	LRG_8:g.24003G>A	NM_006920.4:c.602+1G>A	p.?	Allen *et al.*[11]	EE
CUX2	Missense	Chr12(GRCh37):g.111748354G>A	NM_015267.3:c.1768G>A	p.Glu590Lys	Rauch *et al.*[2]	ID
CUX2	Missense	Chr12(GRCh37):g.111748354G>A	NM_015267.3:c.1768G>A	p.Glu590Lys	Allen *et al.*[11]	EE
SCN2A	Missense	Chr2(GRCh37):g.166198975G>A	NM_021007.2:c.2558G>A	p.Arg853Gln	Allen *et al.*[11]	EE
SCN2A	Missense	Chr2(GRCh37):g.166198975G>A	NM_021007.2:c.2558G>A	p.Arg853Gln	Allen *et al.*[11]	EE
DUSP15	Missense	Chr20(GRCh37):g.30450489G>A	NM_080611.2:c.320C>T	p.Thr107Met	Neale *et al.*[7]	ASD
DUSP15	Missense	Chr20(GRCh37):g.30450489G>A	NM_080611.2:c.320C>T	p.Thr107Met	Fromer *et al.*[10]	SCZ

EE, epileptic encephalopathies; ASD, autism spectrum disorder; ID, intellectual disability; SCZ, schizophrenia.

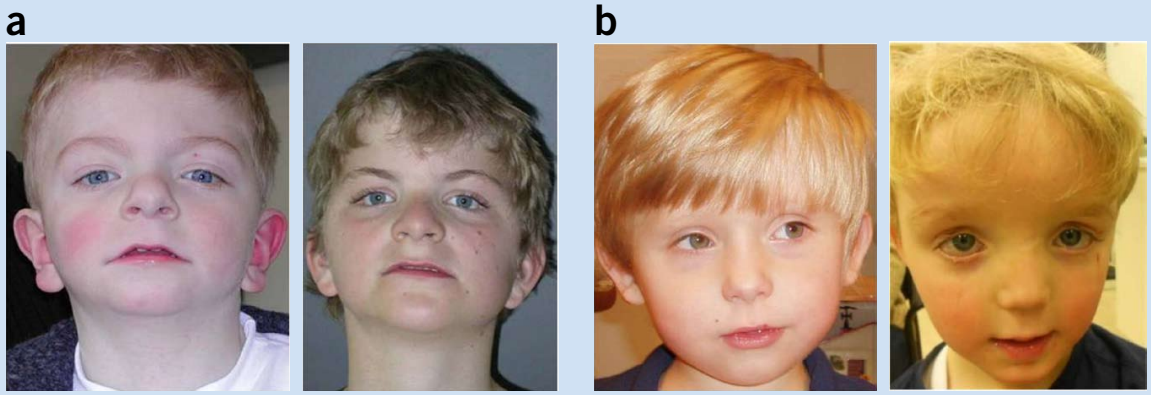
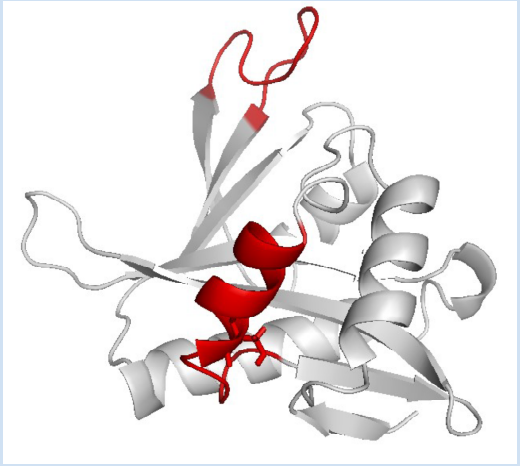


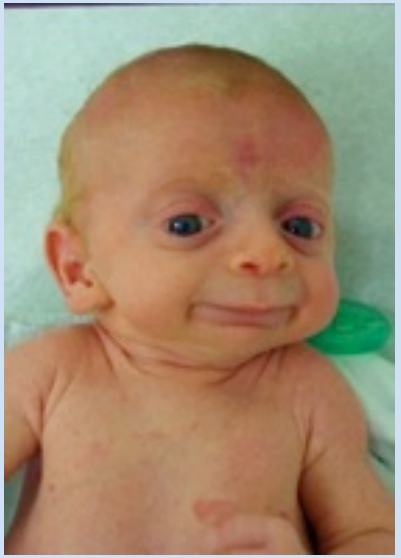
Figure 3 Phenotypic similarity of two patients with identical *PACS1 de novo* mutations and two patients with similar *ADNP* mutations. (**a**) These two unrelated patients show identical *de novo* point mutations (c.607C>T; p.Arg203Trp) in *PACS1* (RefSeq NM_018026.3)[53]. The striking similarity in phenotype includes low anterior hairline, highly arched eyebrows, synophrys, hypertelorism with downslanted palpebral fissures, long eyelashes, a bulbous nasal tip, a flat philtrum with a thin upper lip, downturned corners of the mouth and low-set ears. Reprinted from ref. 53, Copyright (2012), with permission from The American Society of Human Genetics. (**b**) These two unrelated patients both show LoF mutations in *ADNP* (c.2496_2499delTAAA; p.Asp832Lysfs*80 and c.2157C>G; p.Tyr719*)[44] resulting in a new SWI-SNF–related autism syndrome. Patients present with clinical similarities, including a prominent forehead, a thin upper lip and a broad nasal bridge. Reprinted from ref. 44.

Big Questions though:



Simulated structure of S37P mutant

?



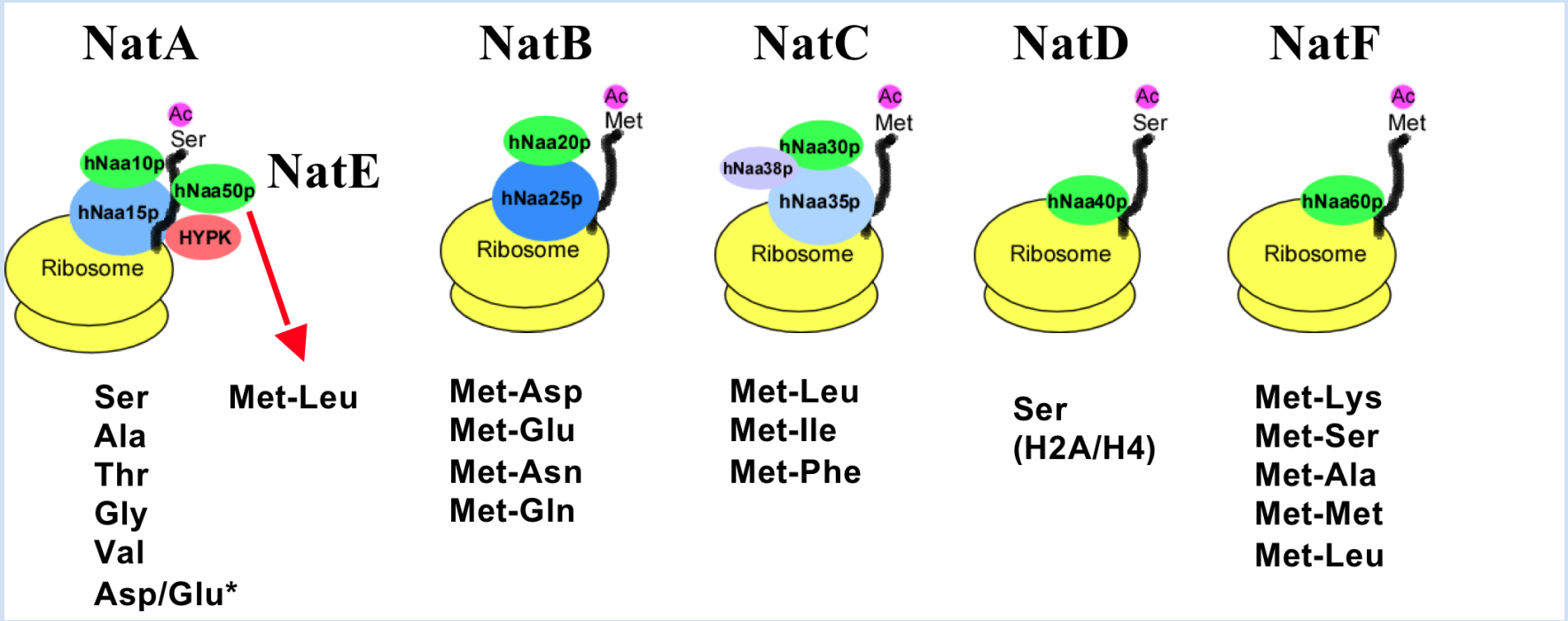
What is the molecular basis of Ogden syndrome?

- Naa10/Naa15 complex
- Naa10 localisation
- Naa10 function

what can we learn from Ogden syndrome?

- characterizing different model systems (fibroblasts, yeast, *C. elegans)*

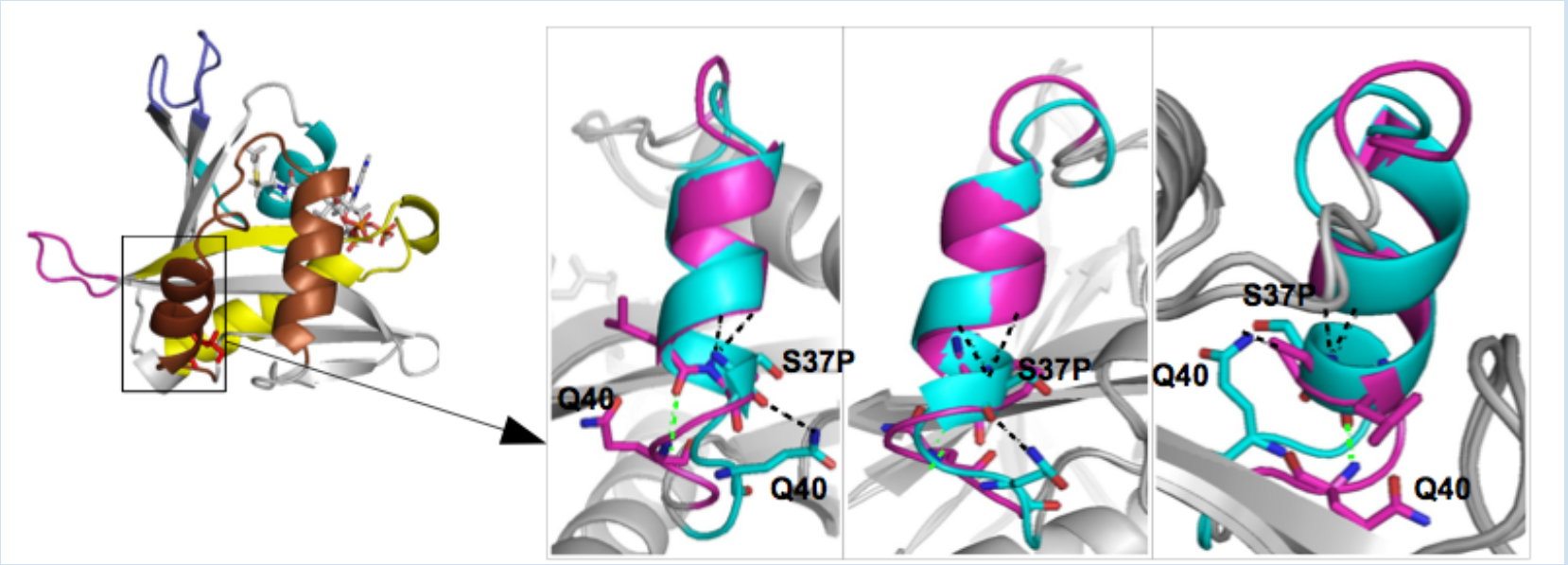
The mutation disrupts the N-terminal acetylation machinery (NatA) in human cells.



Slide courtesy of Thomas Arnesen

The mutation is a missense resulting in Serine to Proline change in Naa10p

- Ser 37 is conserved from yeast to human
- Ser37Pro is predicted to affect functionality (SIFT and other prediction programs)
- Structural modelling of hNaa10p wt (cyan) and S37P (pink)



Open question:
Function of N-terminal acetylation?

Protein stability? Protein secretion?

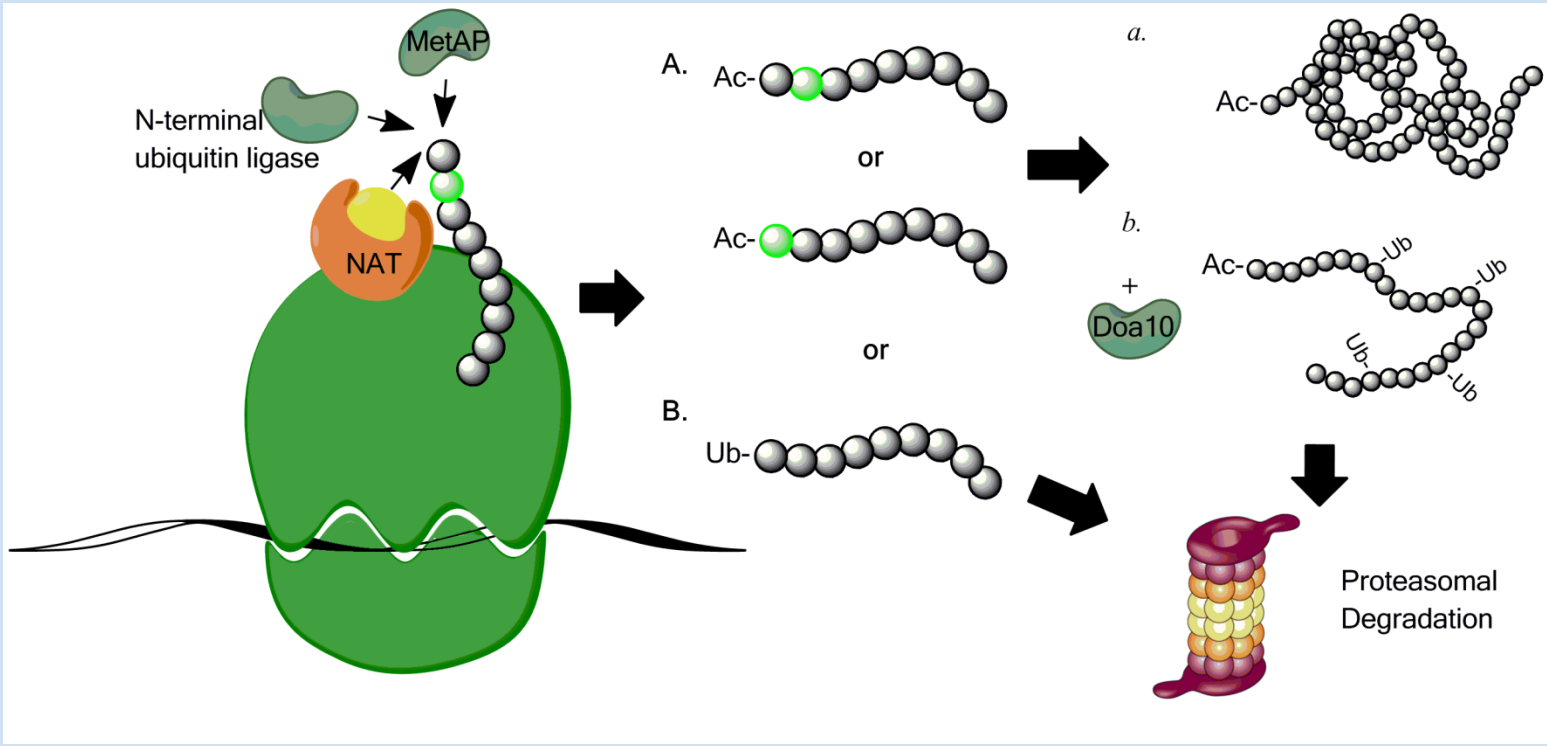
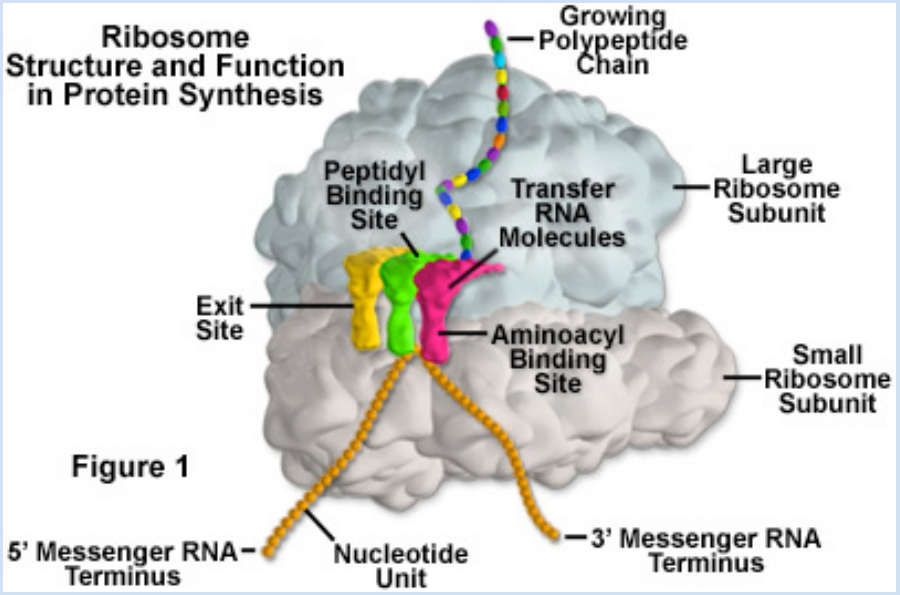
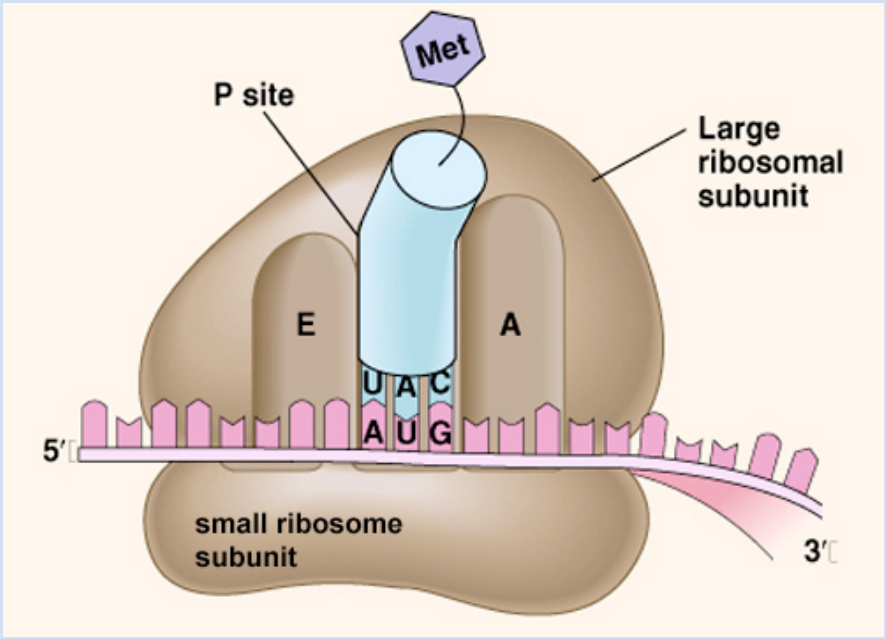
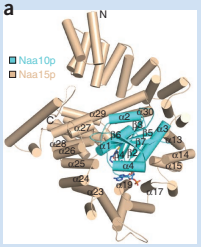


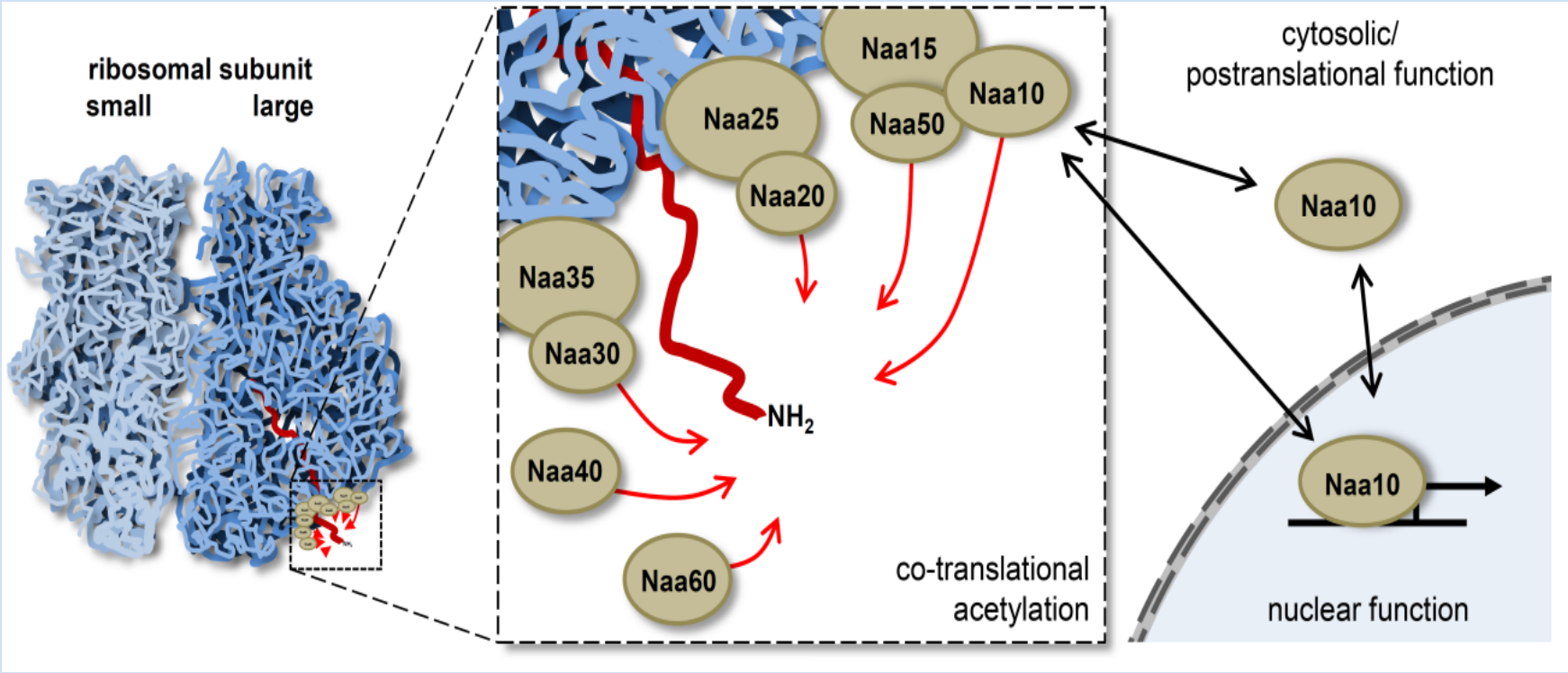
Figure courtesy of Kris Gevaert


Ribosome
Structure and Function
in Protein Synthesis
Growing
Polypeptide
Chain
Peptidyl
Binding
Site
Transfer
RNA
Molecules
Large
Ribosome
Subunit
Exit
Site
Aminoacyl
Binding
Site
Small
Ribosome
Subunit
Figure 1
5' Messenger RNA
Terminus
Nucleotide
Unit
3' Messenger RNA
Terminus


Met
P site
Large
ribosomal
subunit
E
A
U A C
A U G
5'
small ribosome
subunit
3'

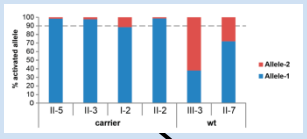

a
N
Naa10p
Naa15p
C

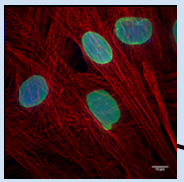
NAA10 + NAA15



- Naa10 co-translationally acetylates the N^{α}-terminal amino group of the nascent polypeptide chains of classical substrates as they emerge from the ribosome.

patient cells

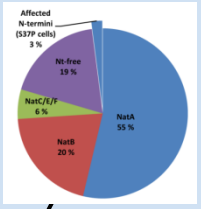




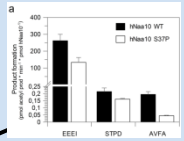
X-chromosome skewing in carriers

normal nuclear envelope

molecular level

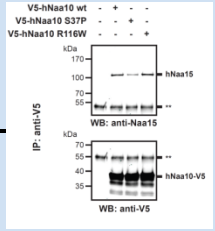


affected *in vivo* substrates



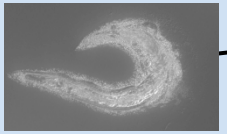
reduced enzymatic activity

Naa10 S37P



disrupted NatA complex

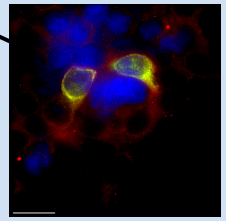
C. elegans



initial experiments

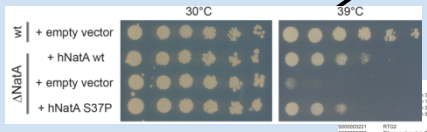
HEK293 cells

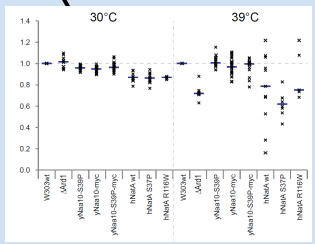
normal localization



growth defect at 39°C
ribosome defect?
no difference cyclo/doxo?

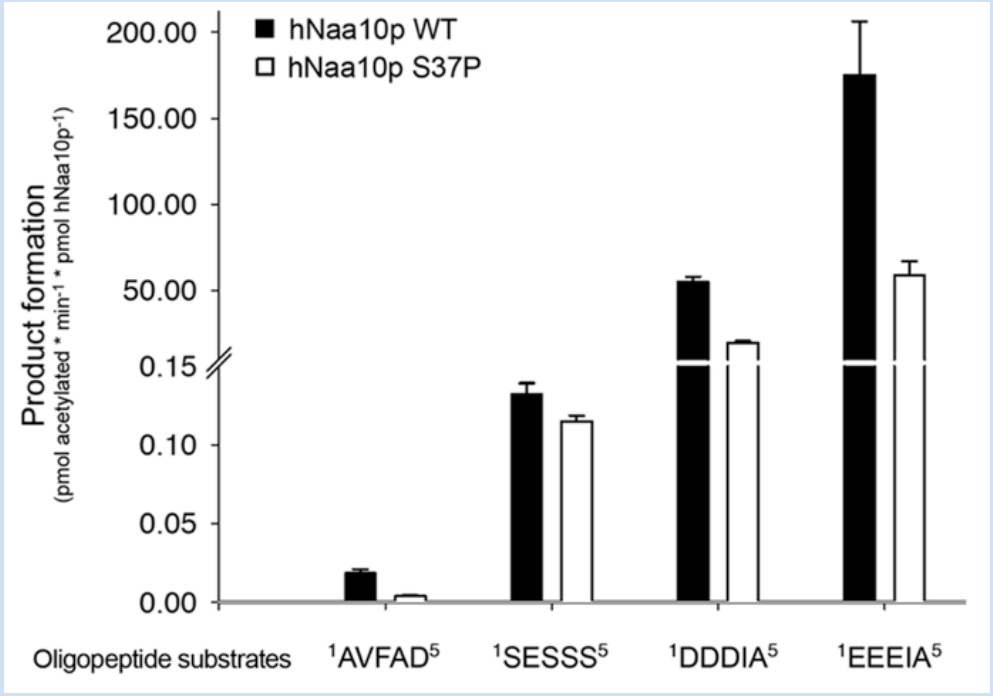
S39 mutation?
growth defect of "humanized" strains

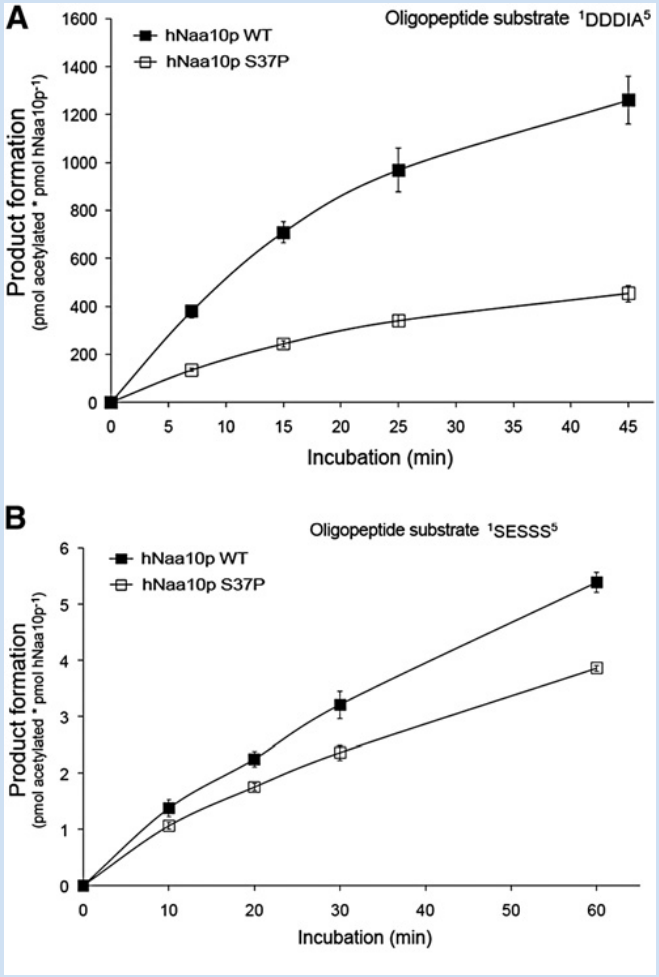




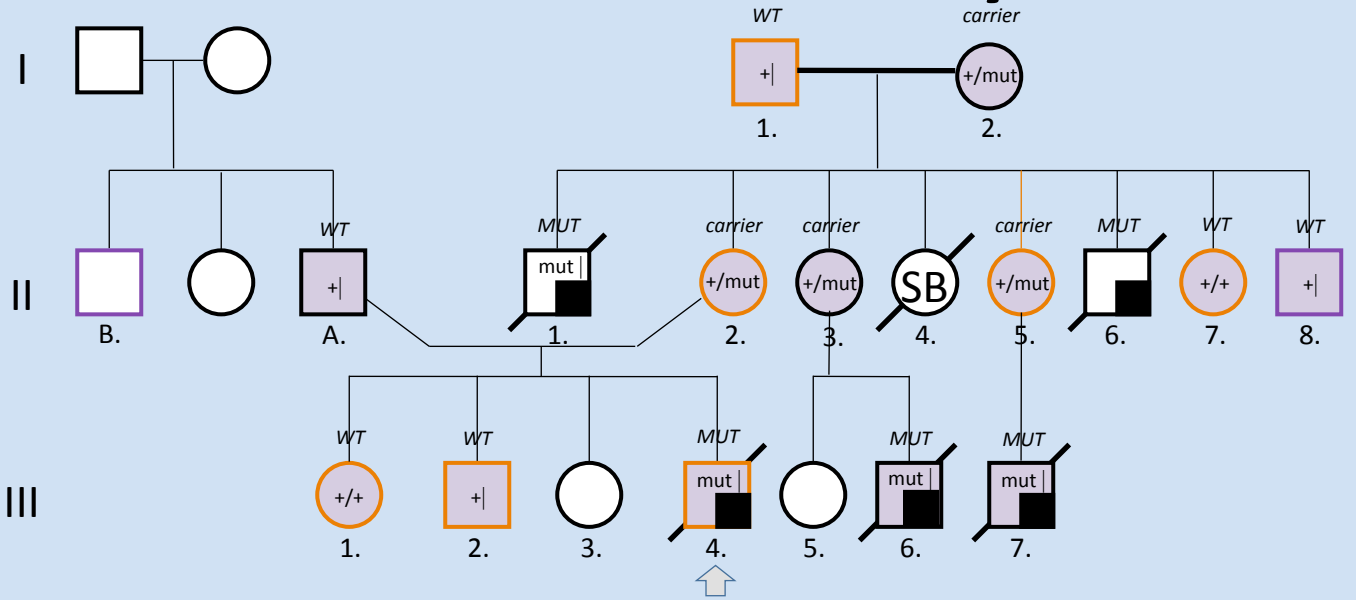
S. cerevisiae

NAT activity of recombinant hNaa10p WT or p.Ser37Pro towards synthetic N-terminal peptides





Proteomics Analysis of EBV-transformed cell lines and fibroblasts from family members



= male ♂

= female ♀

= FFPE DNA (for patient III.7.) or DNA from blood available (and for some of them: EBV transformed cell lines available + skin fibroblast of patient III.6.)

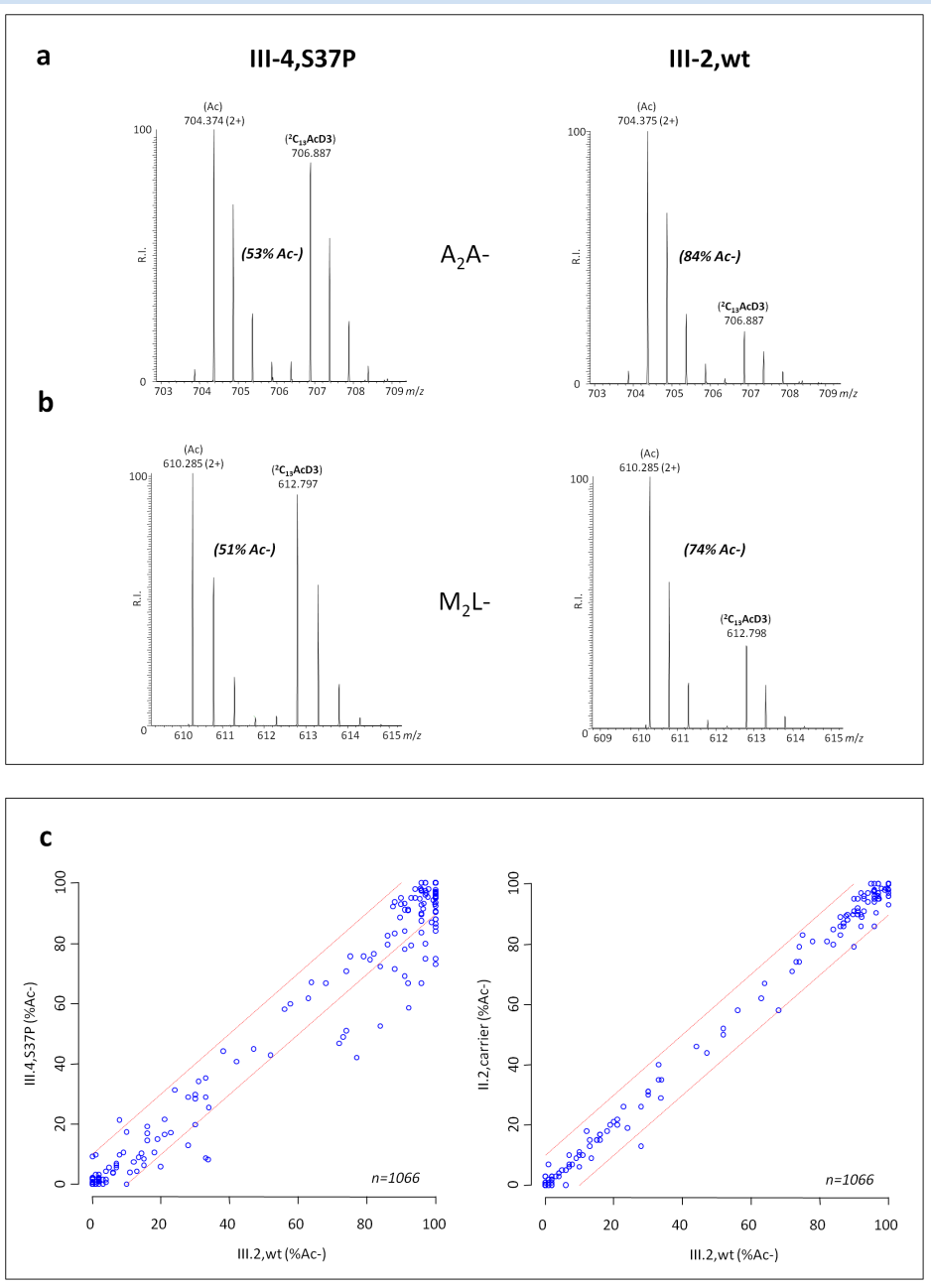
SB = stillborn

= proband

= patient samples analyzed by N-terminal COFRADIC analyses **(#1 to #5)**

= patient samples prepared for N-terminal COFRADIC analyses (but still to be analyzed) **(#8 and #9)**

III.4. proband hemizygous, mutant (89323) **(#1a)** (#1b)
II.2. mother of proband, carrier (89324) **(#2)**
II.A. married-in father of proband, WT(89325)
III.2. brother of proband, WT(90526) **(#3)**
III.1. sister of proband, WT (90527) **(#4)**
I.2. grandmother of proband, carrier (90528)
I.1. married-in grandfather of proband, WT(90529) **(#5)**
II.7. aunt of proband, WT (90530) **(#6)**
II.3. aunt of proband, carrier (90531)
II.B. married-in uncle of proband, WT(90532) (#8)
II.8. uncle of proband, WT(90688) (#9)
II.5. aunt of proband, carrier with deceased boy (90797) **(#7)**


a
III-4,S37P
III-2,wt
(Ac)
704.374 (2+)
(2C13AcD3)
706.887
(53% Ac-)
A2A-
(Ac)
704.375 (2+)
(84% Ac-)
(2C13AcD3)
706.887
R.I.
m/z
b
(Ac)
610.285 (2+)
(2C13AcD3)
612.797
(51% Ac-)
M2L-
(Ac)
610.285 (2+)
(74% Ac-)
(2C13AcD3)
612.798
c
III.4,S37P (%Ac-)
III.2,wt (%Ac-)
n=1066
II.2,carrier (%Ac-)
III.2,wt (%Ac-)
n=1066

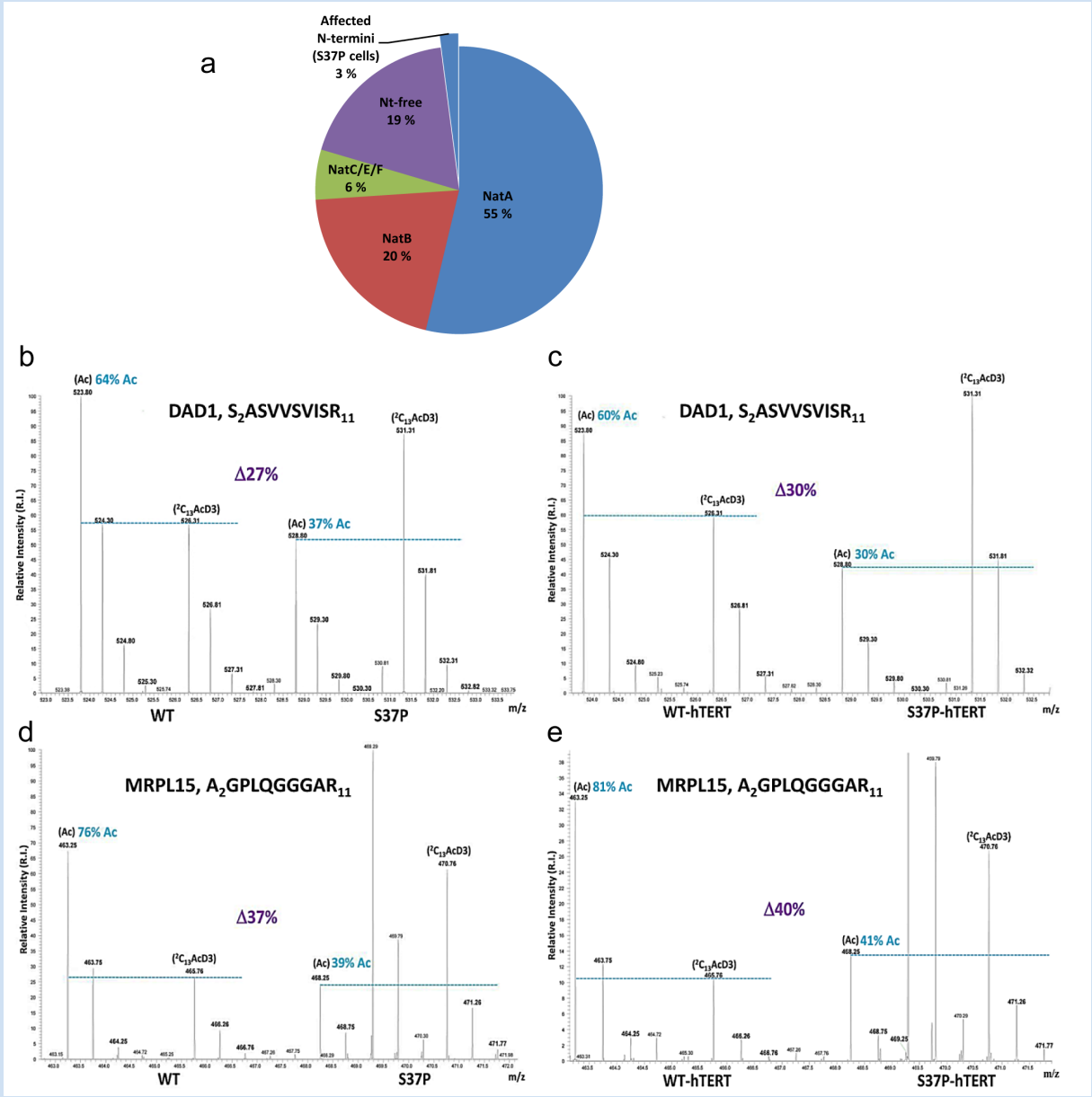

a
Affected N-termini (S37P cells) 3 %
Nt-free 19 %
NatC/E/F 6 %
NatA 55 %
NatB 20 %
b
DAD1, S2ASVVSVISR11
(Ac) 64% Ac
(2C13AcD3)
Δ27%
(Ac) 37% Ac
Relative Intensity (R.I.)
WT
S37P
m/z
c
DAD1, S2ASVVSVISR11
(Ac) 60% Ac
(2C13AcD3)
Δ30%
(Ac) 30% Ac
WT-hTERT
S37P-hTERT
m/z
d
MRPL15, A2GPLQGGGAR11
(Ac) 76% Ac
(2C13AcD3)
Δ37%
(Ac) 39% Ac
WT
S37P
m/z
e
MRPL15, A2GPLQGGGAR11
(Ac) 81% Ac
(2C13AcD3)
Δ40%
(Ac) 41% Ac
WT-hTERT
S37P-hTERT
m/z

Table 1: Overview of N-termini less acetylated in Naa10-S37P B-cells, fibroblasts and siNatA HeLa cells.

NAT type	P1	P1'	P2'	significant B-cells	significant fibroblasts	siNatA (HeLa)	Description
NatA	M	V	N	√	√	√	**Peptidyl-prolyl cis-trans isomerase A**
NatA	M	A	A	√	√	√	**Translational activator GCN1**
NatA	M	A	A	√	√	√	**Transcription elongation factor B polypeptide 3**
NatA	M	A	V	√	√	√	**Ribonuclease P protein subunit p30**
NatA	M	G	A	√	√		**THO complex subunit 7 homolog**
NatA	M	S	A	√	√		**Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit DAD1**
NatA	M	T	M	√	√		**14-3-3 protein beta/alpha**
NatA	M	A	G	√	√		**39S ribosomal protein L15, mitochondrial**
NatA	M	A	A	√	√		**E3 ubiquitin-protein ligase RNF5**
NatA	M	T	K	√	√		**Leucine-rich repeat-containing protein 59**
NatA	M	A	V	√	√		**Ras GTPase-activating protein 3**
other	-	M	V	√	√		**Peptidyl-prolyl cis-trans isomerase A**
other	-	M	V	√	√		**SUMO-activating enzyme subunit 1**
NatA	M	A	E	√		√	**60S ribosomal protein L13a**
NatA	M	V	E	√		√	**SUMO-activating enzyme subunit 1**
NatA	M	A	L	√			**26S protease regulatory subunit 8**
NatA	M	A	Q	√			**Serum response factor-binding protein 1**
NatA	M	V	E	√			**Protein LCHN**
NatA	M	S	G	√			**Transmembrane protein 50A**
NatA	M	T	A	√			**Transmembrane protein 85**
NatC or other	-	M	L	√			**Kinesin-like protein KIF21A**
NatC or other	-	M	L	√			**p53 and DNA damage-regulated protein 1**
other	-	M	V	√			**Deoxyhypusine hydroxylase**
other	-	M	M	√			**Uncharacterized protein C11orf46**
NatA	M	A	A	√	√		**Epidermal growth factor receptor substrate 15**
NatA	M	S	T		√		**Mediator of RNA polymerase II transcription subunit 30**
NatA	M	A	A		√		**AN1-type zinc finger protein 5**
NatA	M	G	A		√		RNA-binding protein 7

Summary

- Found first human genetic disease with proof involving Nt-acetylation of proteins.
- Characterizing the Nt-acetylation pathway both *in vitro* and *in vivo* to reveal fundamental new biology, including possible drug targets.
- Identifying interacting components and more downstream substrates of NatA complex.

The End