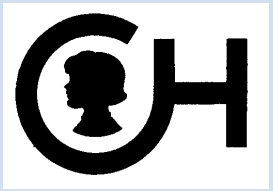
Finding and Analyzing Human Genetic Variation in Neuropsychiatric Disorders

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



The Big Picture

- Over the course of my entire career, I want to help understand the pathophysiology of severe mental disorders, including such things as developmental delay, mental retardation, psychotic disorders (schizophrenia, bipolar, schizoaffective), Tourette Syndrome and obsessive compulsive disorder.

- I expect this to uncover new biology along the way.

The toll with Brain Disorders is tremendous.

Most recent analysis in Europe showed that brain disorders cost almost US $1 trillion year per year, more than cancer, cardiovascular disease and diabetes combined.

These brain disorders include:
Mood disorders
Psychotic Disorders
Addiction
Anxiety
Dementia
Headache
Other- brain tumor, **child/adolescent developmental disorders (autism, ADHD, tics, etc…),** eating disorders, epilepsy, **mental retardation,** multiple sclerosis, neuromuscular disorders, Parkinson's, personality disorders, sleep disorders, somatoform disorder, stroke and traumatic brain injury.

* Cost of disorders of the brain in Europe 2010.
Gustavsson A, et al.
Eur Neuropsychopharmacol. 2011 Oct;21(10):718-779. Epub 2011 Sep 15.

Scientific Basis for the Support of Biomedical Science

Julius H. Comroe, Jr., and Robert D. Dripps

Science. 1976 Apr 9;192(4235):105-11.

"scientific advance requires far more work than that reported by the discoverer or by those who wrote key articles essential for his discovery…. scientists earlier and later than the discoverer have always been essential to each discovery and its full development."

The Story began for me at least by 1993….

when I joined the lab of Don St.Germain to study the role of thyroid hormone in cretinism, which is caused by lack of iodine during maternal pregnancy, so this is an environmentally triggered disease.

Slater SD (2010). The discovery of thyroid replacement therapy.

© Dr Stefan Slater, 80 Whitehouse Road, Cramond, Edinburgh EH4 6PD, Scotland.

cyanosed; the skin of the palms is wrinkled and dry. The legs are short and stout, and the calves are very large, hard,

FIG. 20.—Sporadic cretinism (Case 4), showing the appearance of the patient before treatment. The myxœdematous swelling of the face, eyelids, lips, tongue, neck, abdomen, forearms, hands, thighs, and calves, is admirably shown. (Contrast with Fig. 21.)

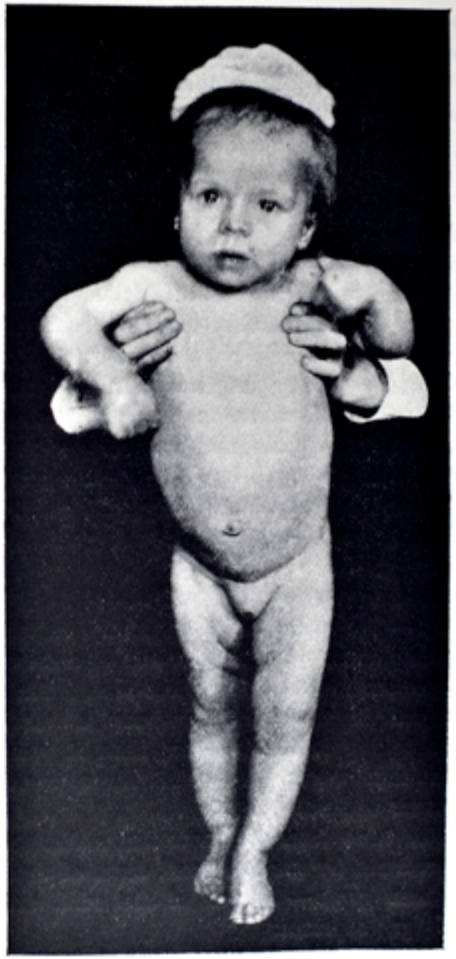
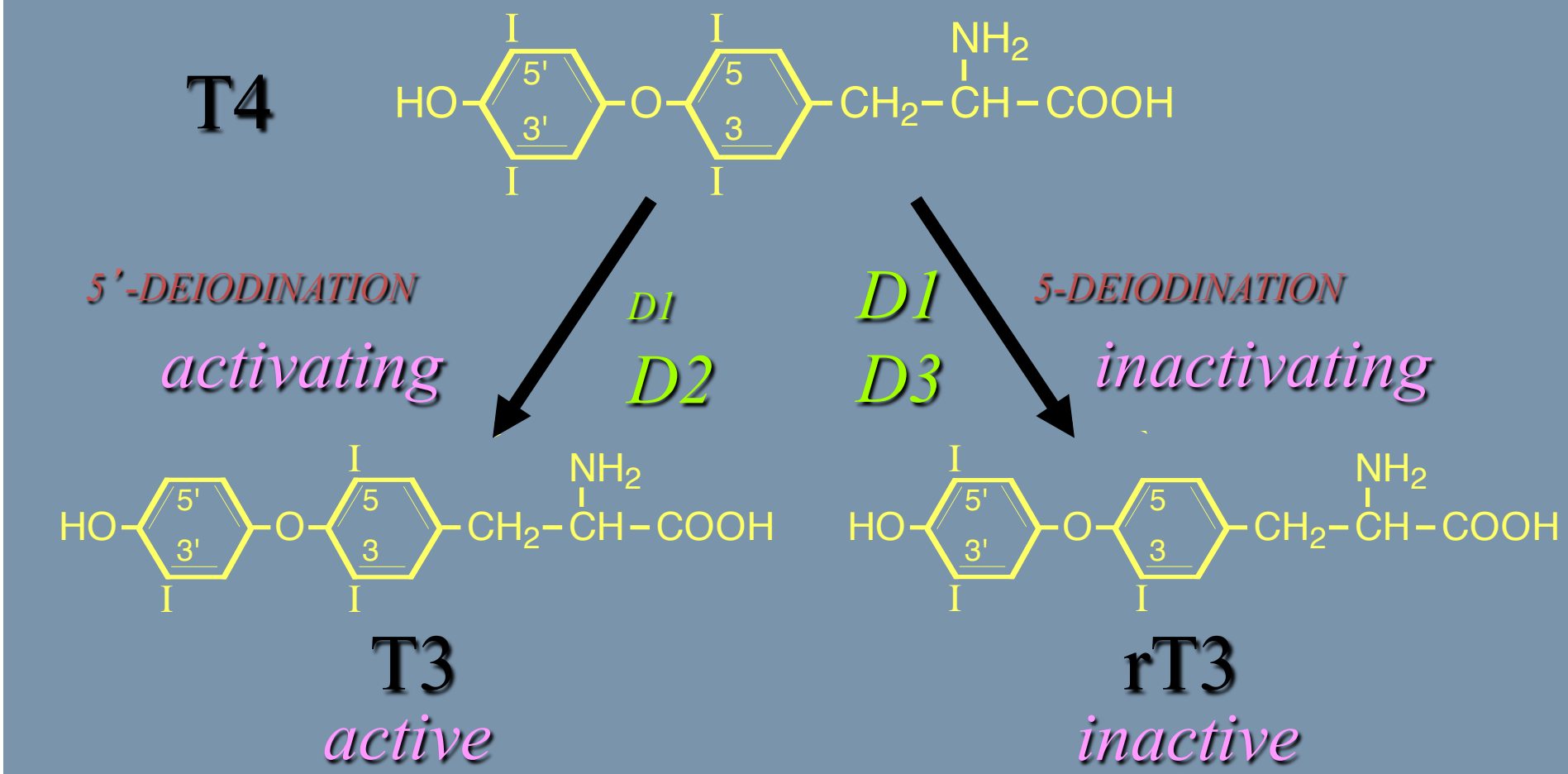


FIG. 21.—Sporadic cretinism (Case 4), showing the appearance of the patient two months after the commencement of the thyroid treatment. Figs. 20, 21, 22, and 23 have all been *reproduced exactly to scale*, so that the exact (relative) increase in height, and diminution in breadth and girth, before and after the treatment, are shown. The myxœdematous swelling of the eyelids, lips, tongue, and body generally has disappeared. The tongue is no longer protruded. The abdomen

Thyroid Hormone and Deiodination Pathways



Isolation and Characterization of the Mouse Gene for the Type 3 Iodothyronine Deiodinase*

ARTURO HERNÁNDEZ†, GHOLSON J. LYON‡, MARK J. SCHNEIDER, AND DONALD L. ST. GERMAIN

Departments of Medicine and Physiology, Dartmouth Medical School, Lebanon, New Hampshire 03756

Endocrinology. 1999 Jan;140(1):124-30.

Type 3 deiodinase is critical for the maturation and function of the thyroid axis

Arturo Hernandez,[1] M. Elena Martinez,[1] Steven Fiering,[2] Valerie Anne Galton,[3] and Donald St. Germain[1,3]

[1]Department of Medicine, [2]Department of Microbiology and Immunology, and [3]Department of Physiology, Dartmouth Medical School, Lebanon, New Hampshire, USA.

J Clin Invest. 2006 Feb;116(2):476-84.

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.

*Gilmour, D. T., **Lyon, G. J.,** …., Evans, M. J., and Colledge, W. H. (1998). Mice deficient for the secreted glycoprotein SPARC/osteonectin/BM40 develop normally but show severe age-onset cataract formation and disruption of the lens. Embo J 17(7): 1860-70.

Archibald Garrod

Garrod, Archibald E. 1902. The Incidence of Alkaptonuria: A Study in Chemical Individuality. *Lancet*, vol. ii, pp. 1616–1620.

THE INCIDENCE OF ALKAPTONURIA: A STUDY IN CHEMICAL INDIVIDUALITY

ARCHIBALD E. GARROD

Physician to the Hospital for Sick Children, Great Ormondstreet, Demonstrator of Chemical Pathology at St. Bartholemew's Hospital

ALL THE MORE RECENT WORK on alkaptonuria has tended to show

"Garrod was more of a scientist than a physician. His bedside manner was said to be limited to his interest in his patients' urine samples."

From http://www.dnaftb.org/13/bio.html

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."

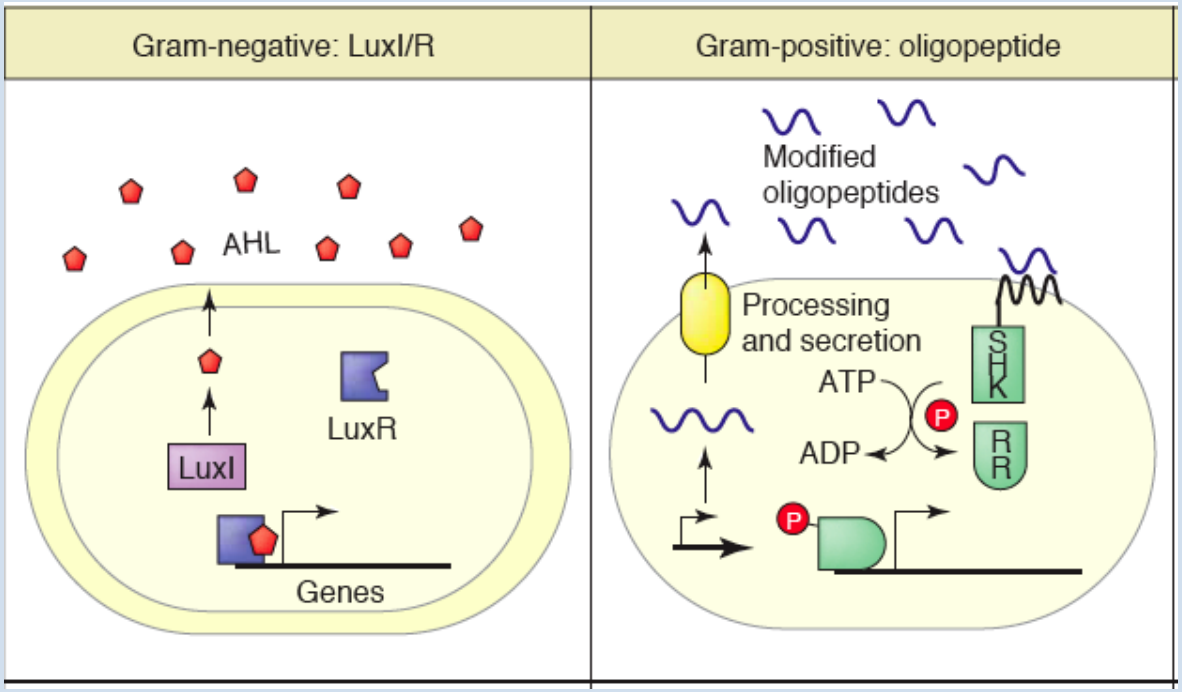
-Olaf Andersen, M.D.
Director of Cornell/Rockefeller/Sloan-Kettering
M.D./Ph.D. program

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

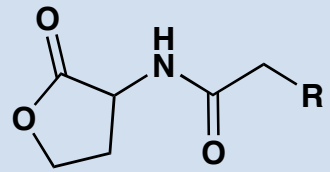
- Much learning of human physiology, anatomy and disease at Weill Cornell Medical College.

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

Bacterial communication via small molecules and peptides



Acyl Homoserine Lactones (AHL's) =



R = $(CH_2)_nCH_3$, $CO(CH_2)_nCH_3$

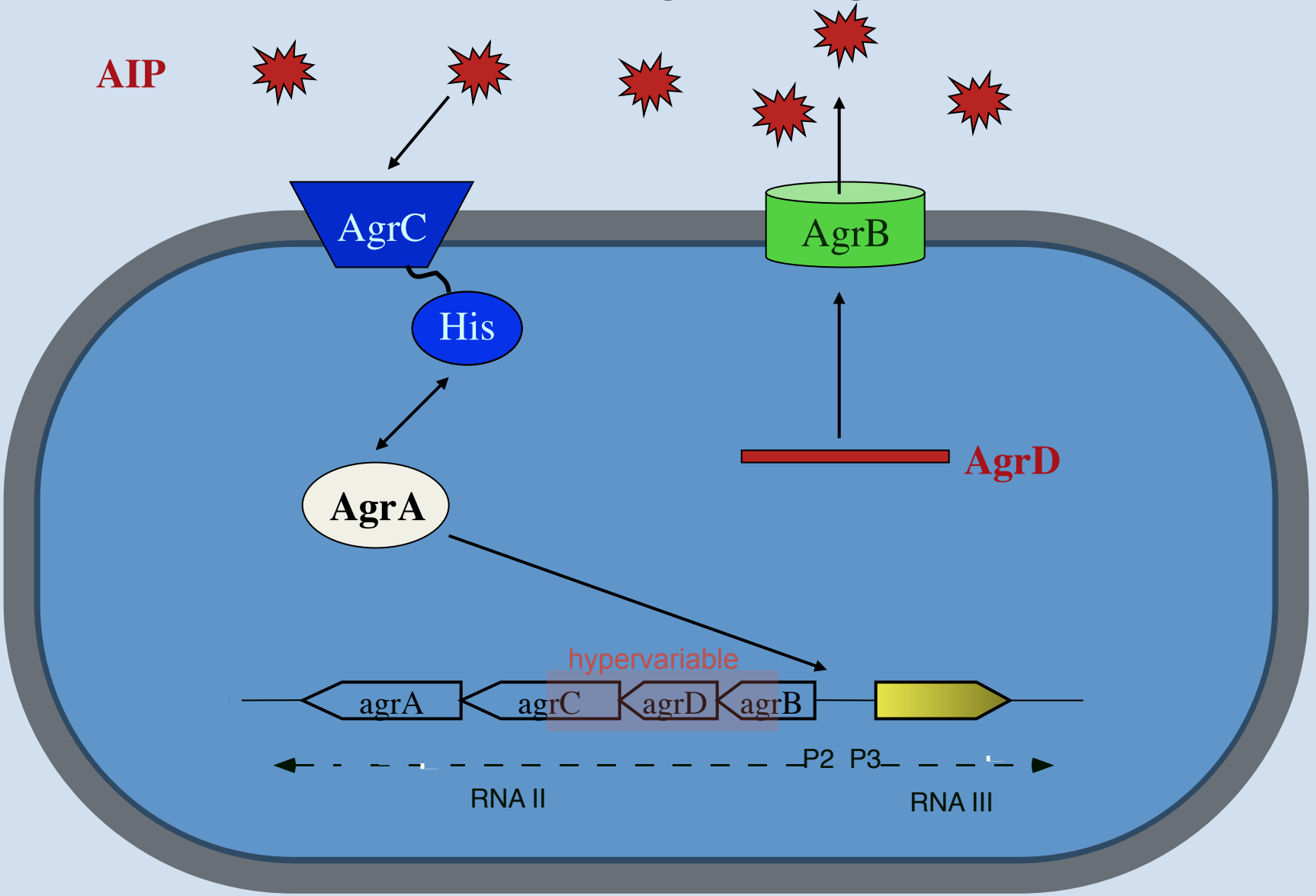
Autoinducing Peptides (AIP's) =

~6-10 amino acid peptides

Modifications include geranylation, thiolactone formation (cyclization)

Lyon, G. & Muir, T.W., *Chem. Biol.* 2004
Bassler, B., *Cell*, 2004, **109**, 421

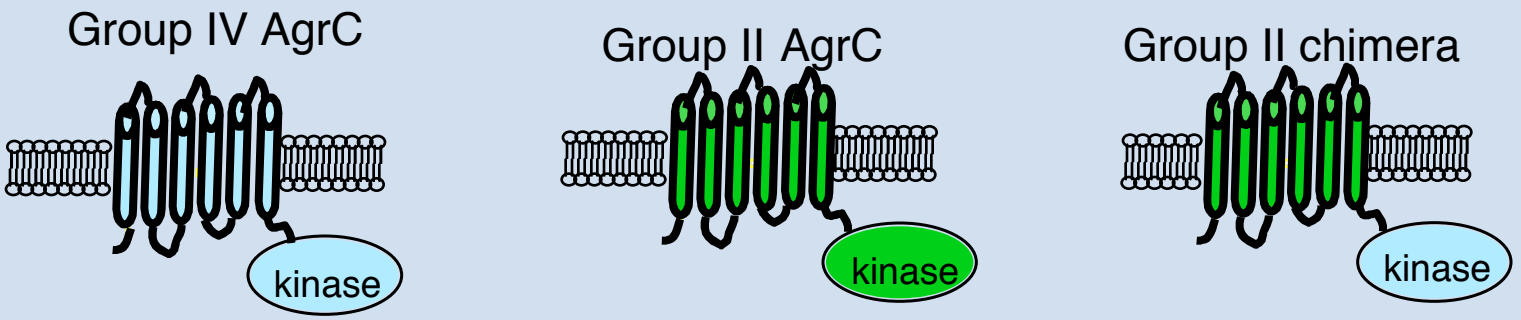
Virulence in Staphylococci is Controlled by a Global Regulon, *agr*



Lyon, G.J., and Novick, R.P. *Peptides,* 2004

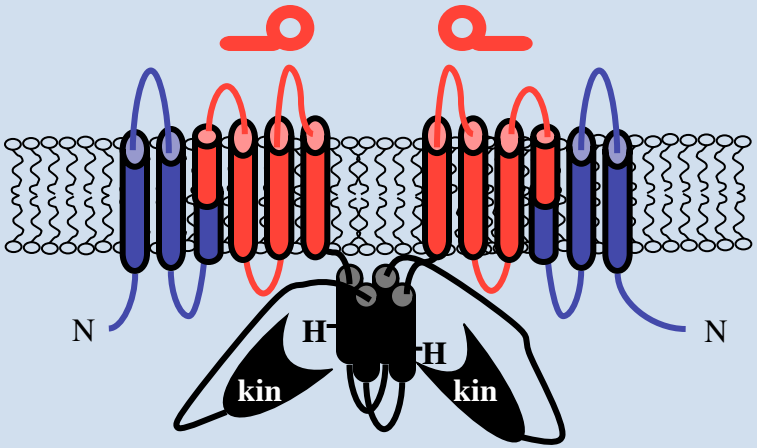
Showing Competitive Antagonism at RHK

1. AIPs Bind Competitively to the AgrC Sensor Domain



Use of sensor/kinase chimeras localize the site of AIP binding; Lyon et al, *JBC,* **2002**

2. AIPs Bind to the Distal Region of the Sensor Domain



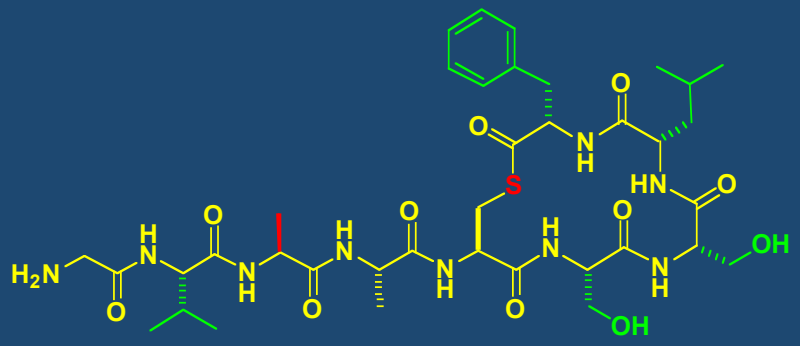
AIP-1 & AIP-IV differ by 1 residue

YSTCDFIM
YSTCYFIM

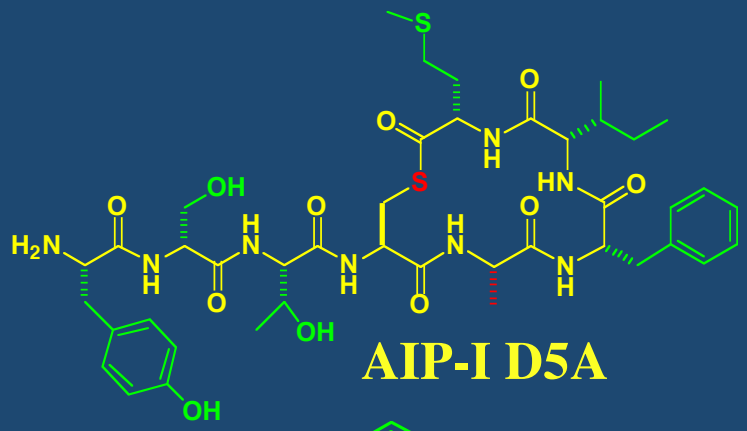
Use of sensor-domain chimeras allow the major binding determinant for group I & IV to be further refined

Wright et al, *PNAS*, 2004

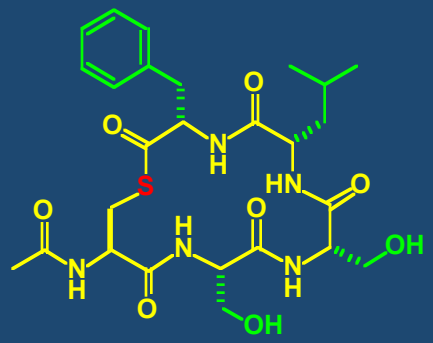
Global Inhibitors of Virulence in *S. aureus*



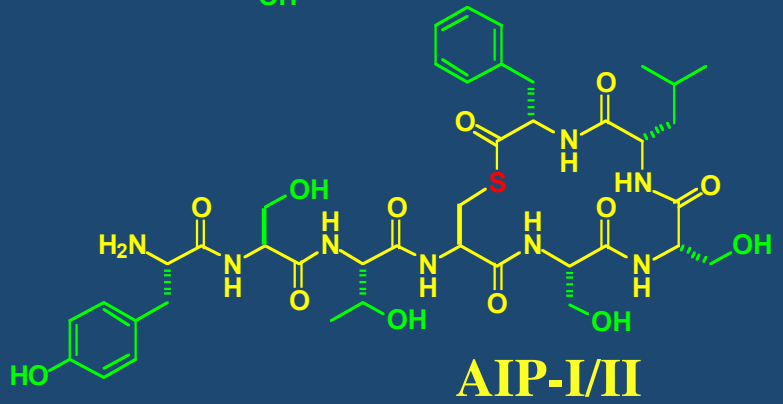
AIP-II N3A



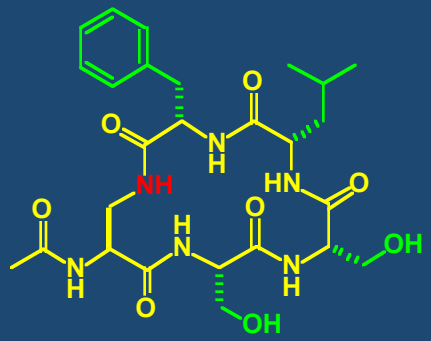
AIP-I D5A



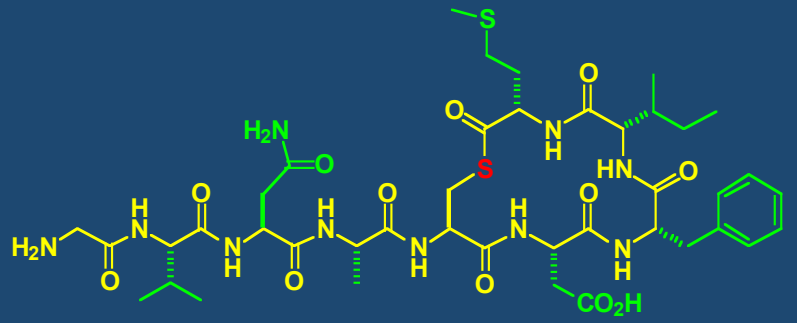
trAIP-II



AIP-I/II



trAIP-II lactam



AIP-II/I

2004- present

Becoming a child, adolescent and adult psychiatrist through clinical residency and practice.

Becoming a scientist through focusing on the genetic basis of neuropsychiatric diseases.

THE UNIVERSITY OF UTAH

UTAH Foundation for BIOMEDICAL research

I moved to Utah in July 2009 to find at least one new human disease, thus revealing new biology.

- **July 2009-December 2009: Attended weekly genetics case conference in which 10-30 genetic cases are presented weekly, led by Dr. Alan Rope and attended by Drs. John Carey and John Opitz.**
- **There are indeed MANY idiopathic disorders not described in the literature, many of which have neuropsychiatric manifestations. I thought about hundreds of such cases, looking for the ideal first family to sequence.**

Discovering a new syndrome and its genetic basis.

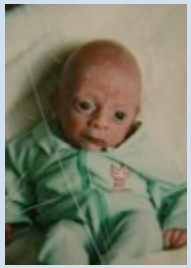
ARTICLE

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

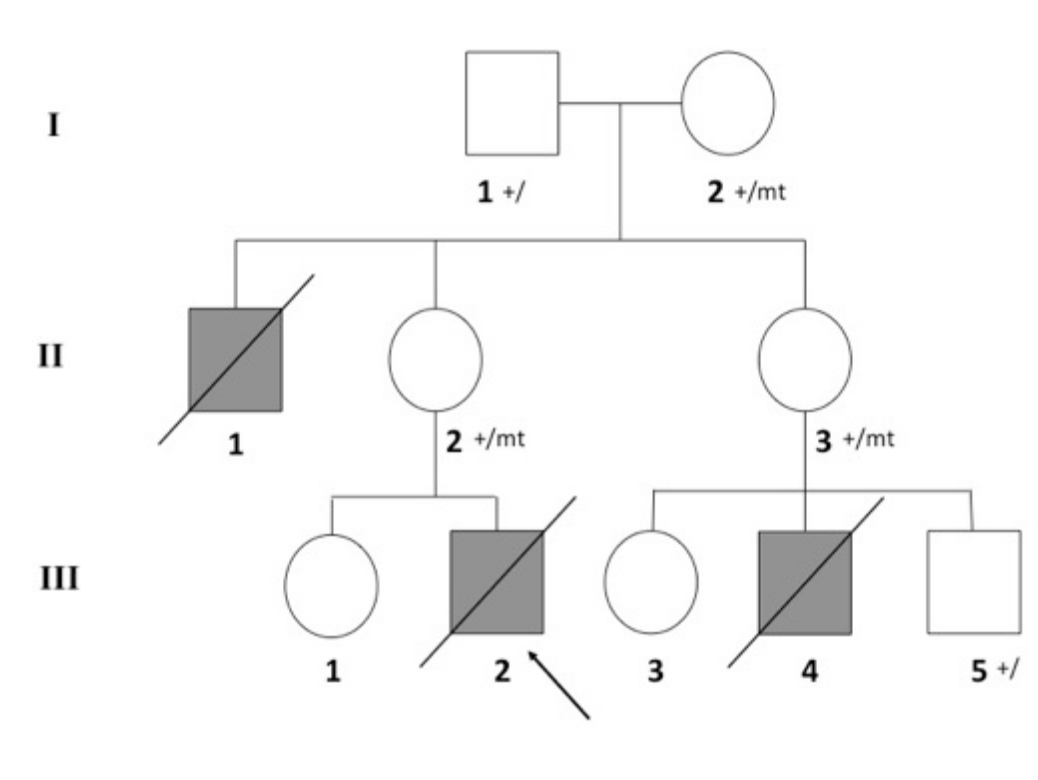
Alan F. Rope,[1] Kai Wang,[2,19] Rune Evjenth,[3] Jinchuan Xing,[4] Jennifer J. Johnston,[5] Jeffrey J. Swensen,[6,7] W. Evan Johnson,[8] Barry Moore,[4] Chad D. Huff,[4] Lynne M. Bird,[9] John C. Carey,[1] John M. Opitz,[1,4,6,10,11] Cathy A. Stevens,[12] Tao Jiang,[13,14] Christa Schank,[8] Heidi Deborah 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,*]

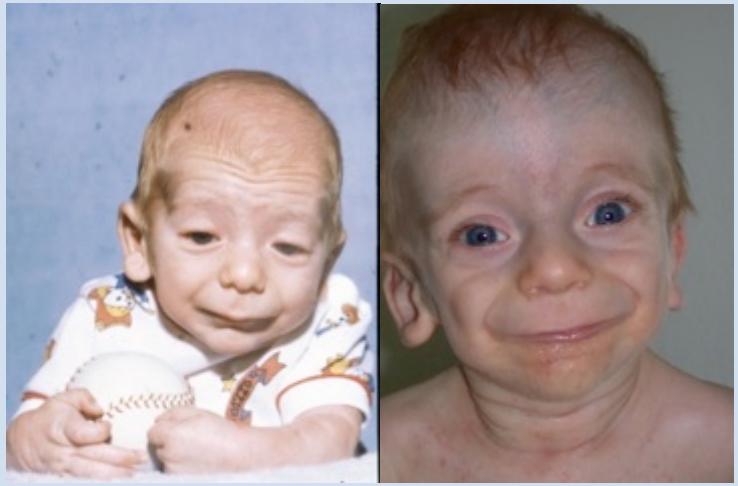
The American Journal of Human Genetics *89*, 1–16, July 15, 2011

Family now in October 2011, with five mutation-positive boys dying from the disease.



An unrelated second family was also identified, due to sharing the same genotype, i.e. the same mutation.

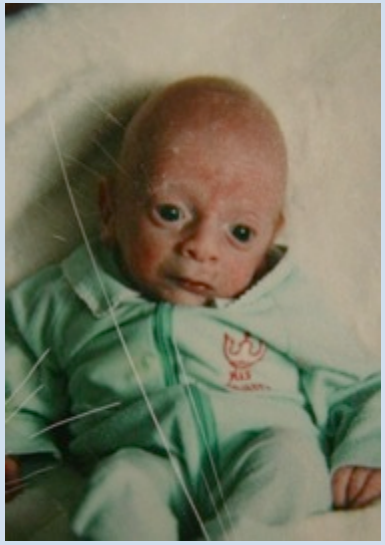
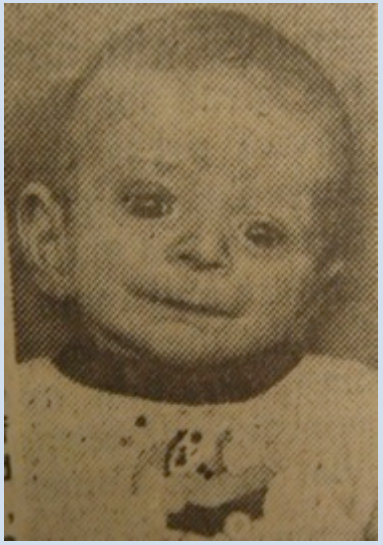
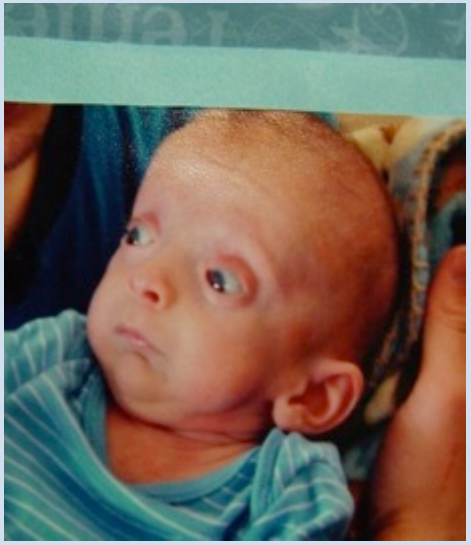
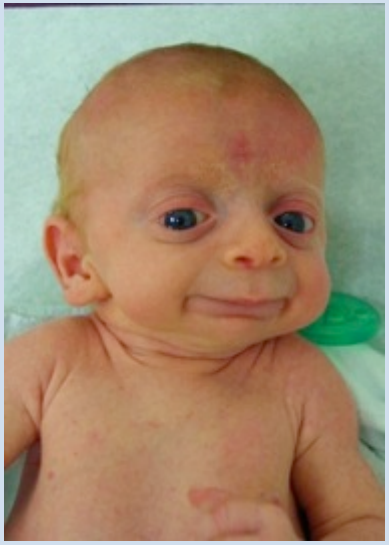
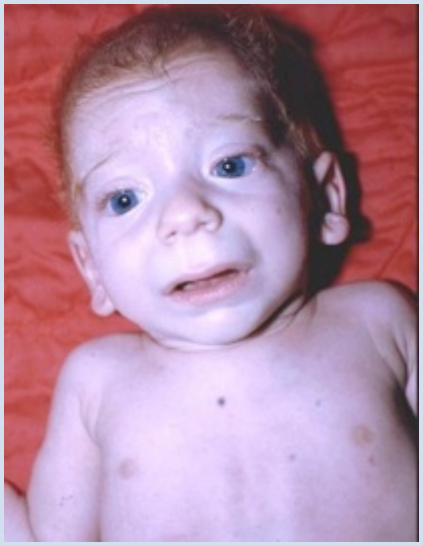
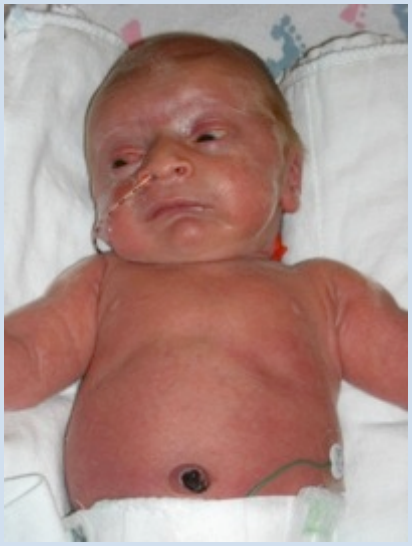
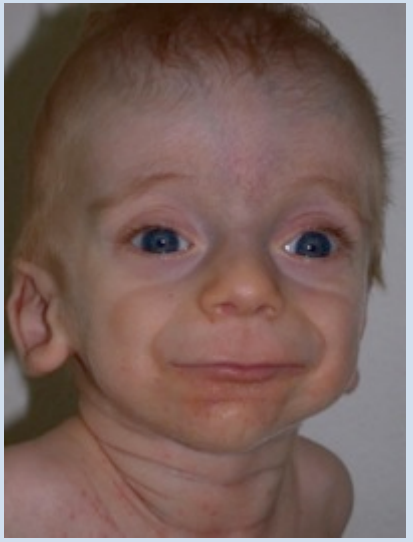




II-1 III-2

Contributed by Les Biesecker and colleagues at NIH

Tentative name: Ogden Syndrome, in honor of where the first family lives, in Ogden, Utah

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.

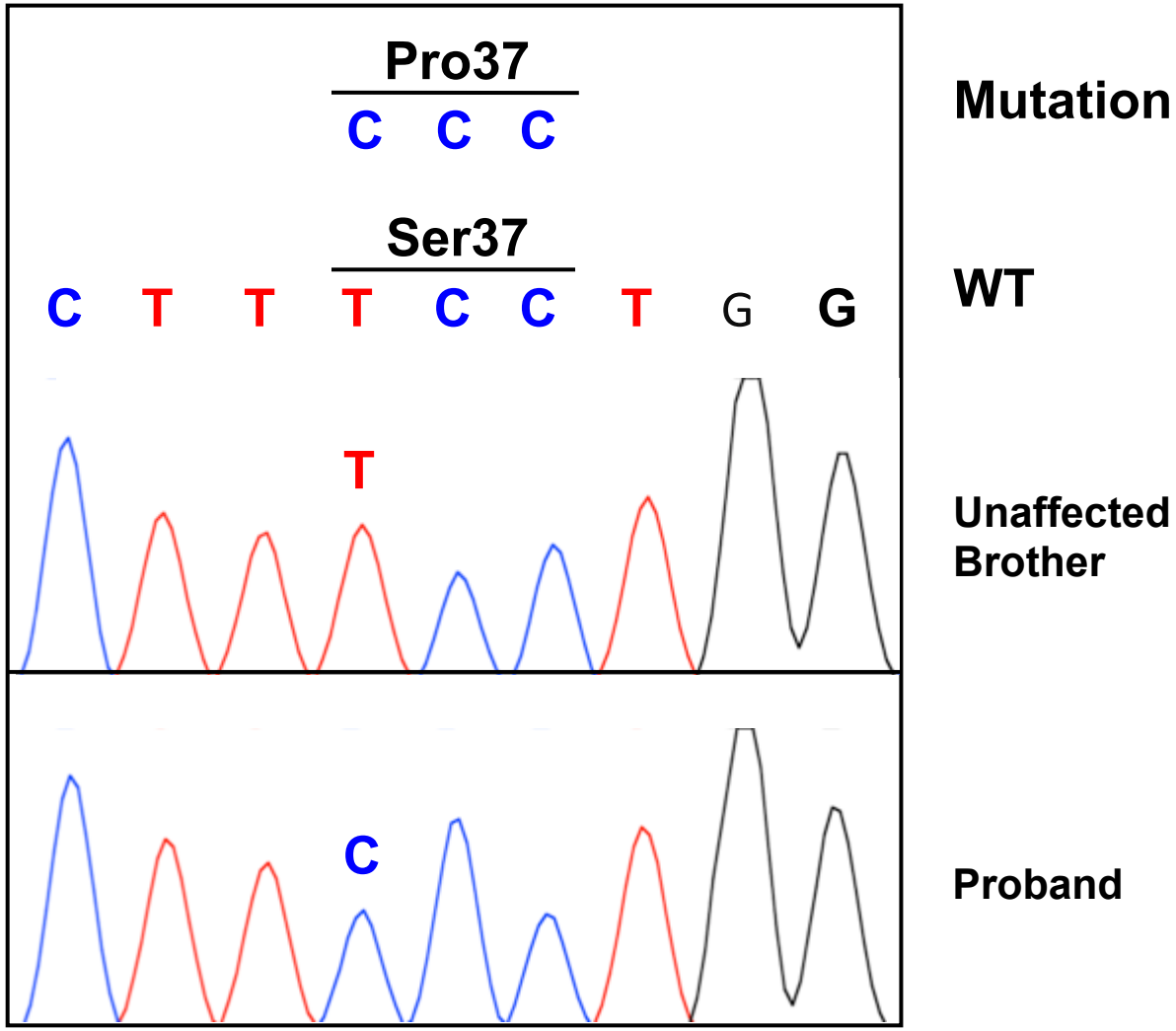
Omicia™ Unlocking Individualized Medicine

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…

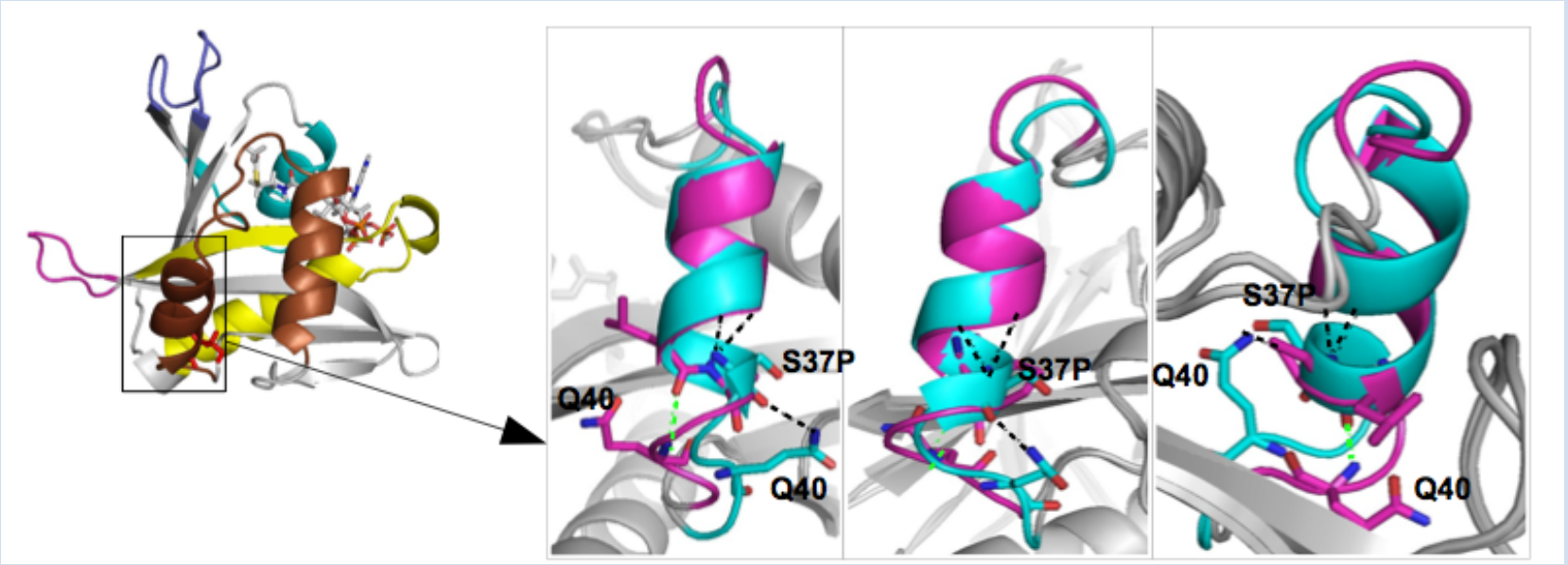


Proving Causality of the mutation

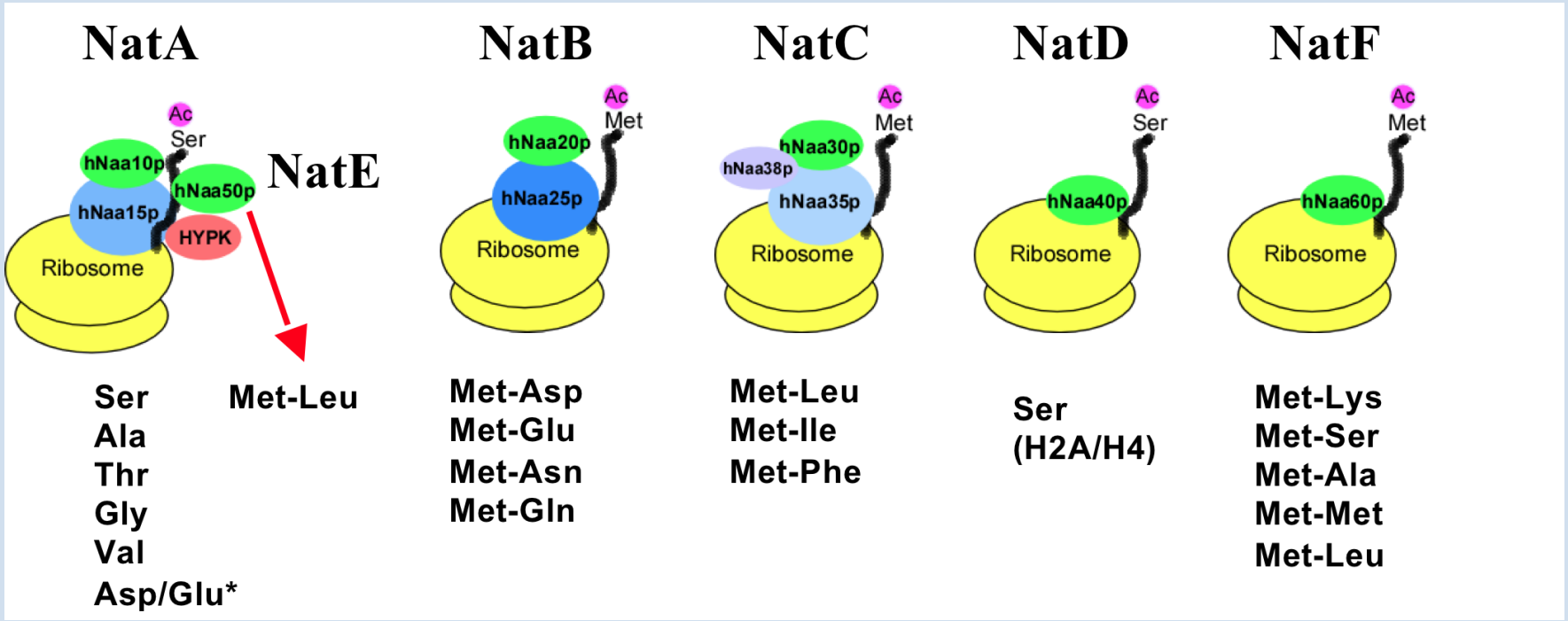
- Present in two unrelated families with very similar phenotype of affected boys.
- Blinded Sanger sequencing showed perfect segregation of the mutation with the disease. Mutation present in Proband, Carrier Mother, Carrier Grandmother and other carrier mothers. Absent in unaffected brother and unaffected uncle.
- Also present in DNA from formalin-fixed paraffin-embedded tissue from two other deceased affected boys, found in pathology department, saved in one case for 30 years.
- Mutation NOT present in ~6000 exomes or genomes sequenced at BGI, CHOP and Utah for other projects.

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)

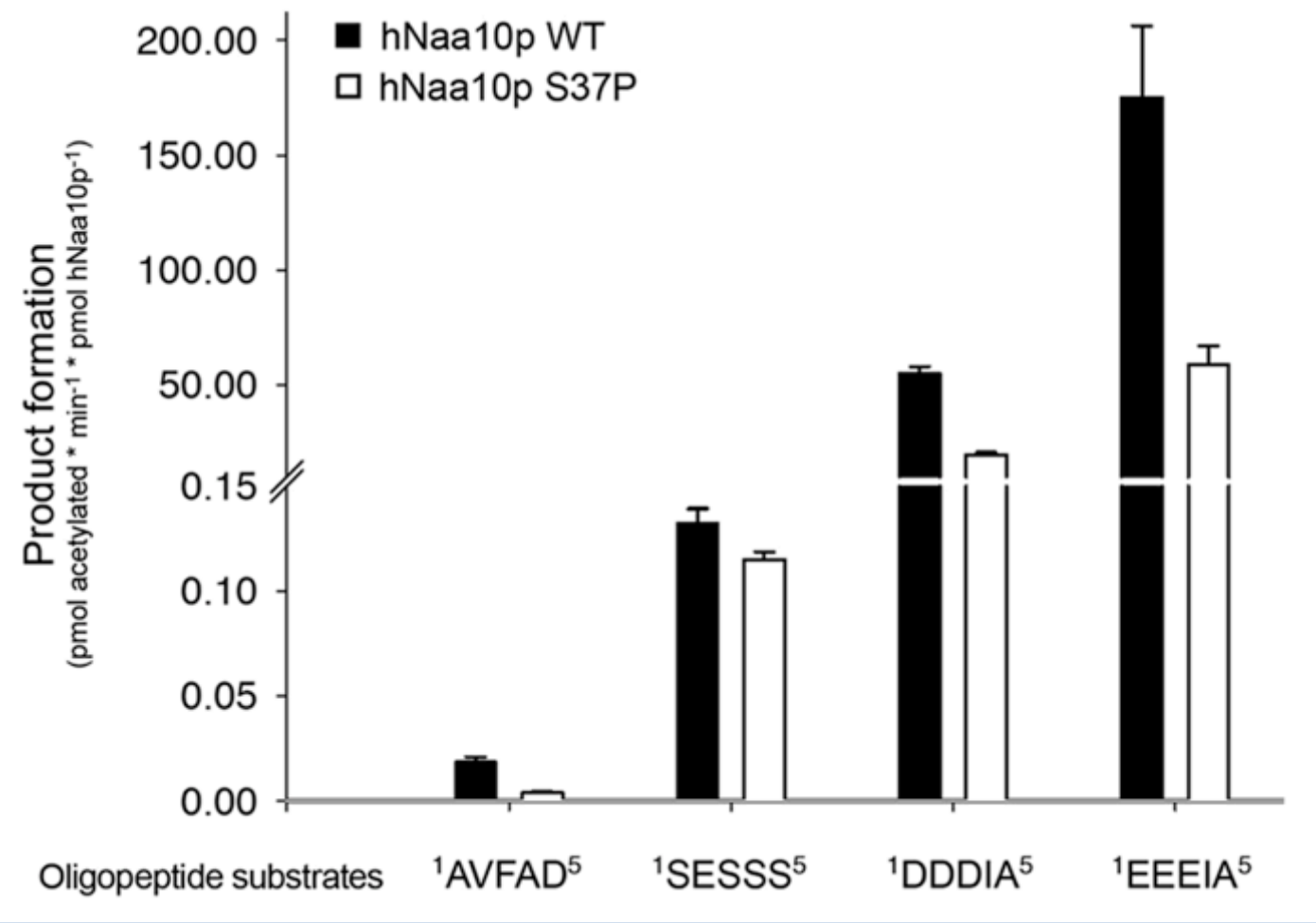


The mutation disrupts the N-terminal acetylation machinery (NatA) in human cells.

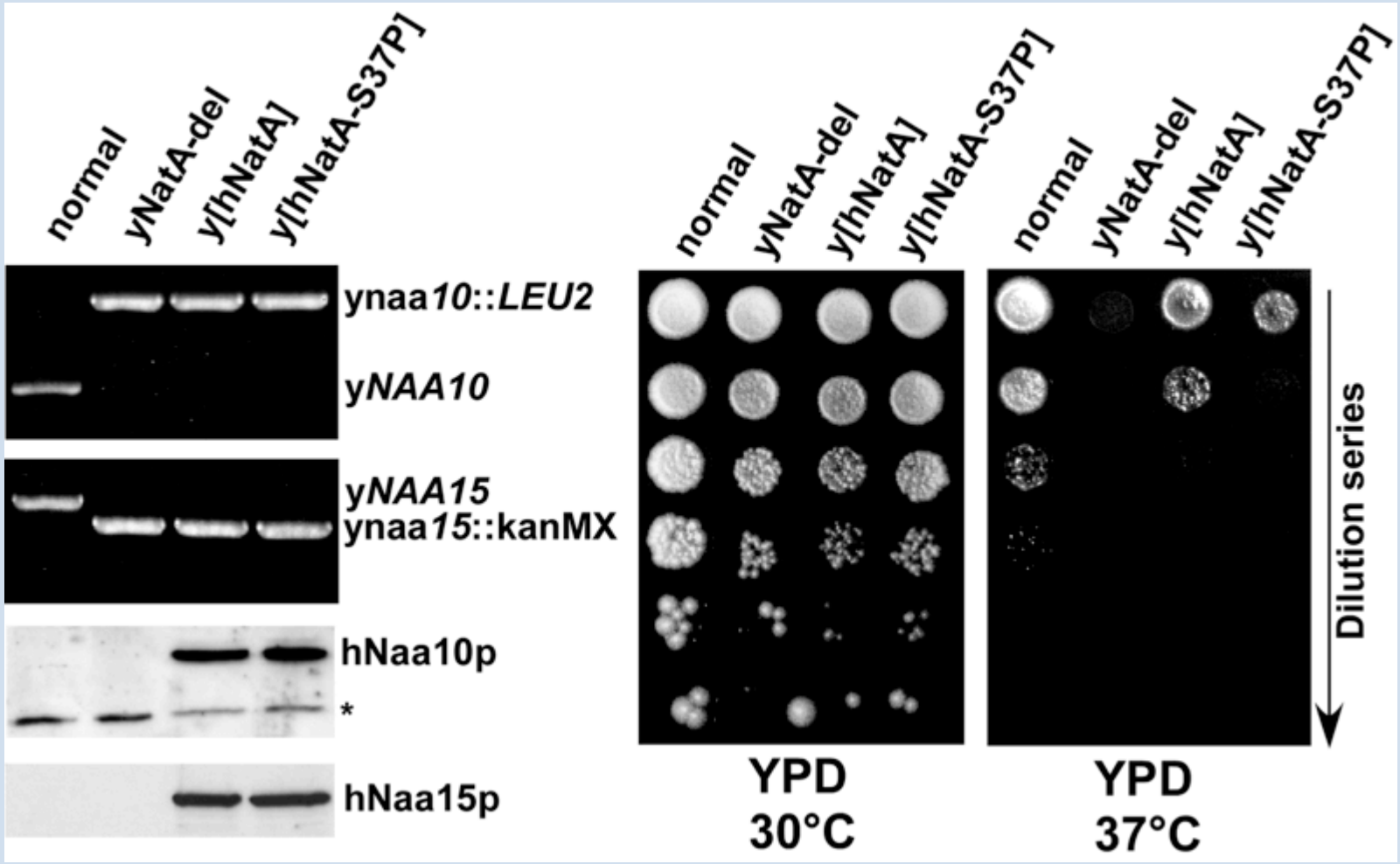


Slide courtesy of Thomas Arnesen

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

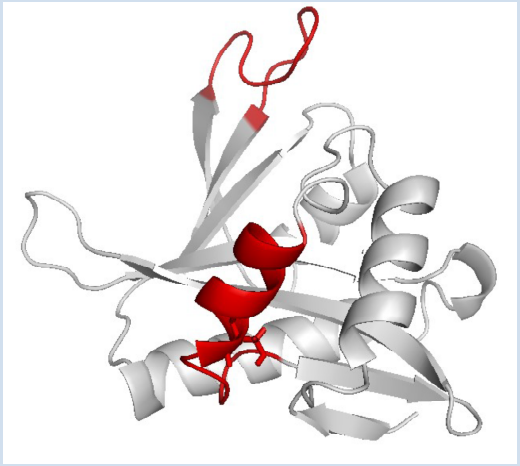


hNaa10p-S37P is functionally impaired *in vivo* using a yeast model.

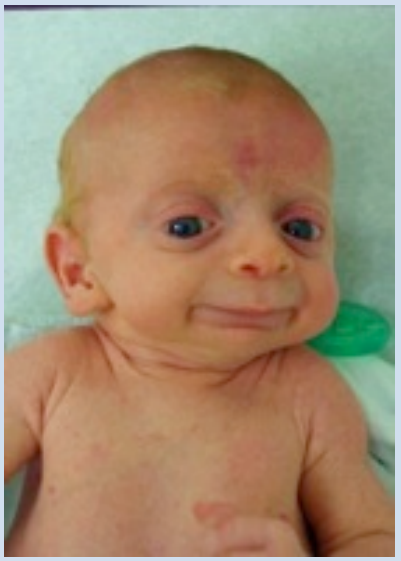


Unpublished data, do not further distribute.

Big Question though:



?



Simulated structure of S37P mutant

Open question:
Function of N-terminal acetylation?

Protein stability? Protein secretion?

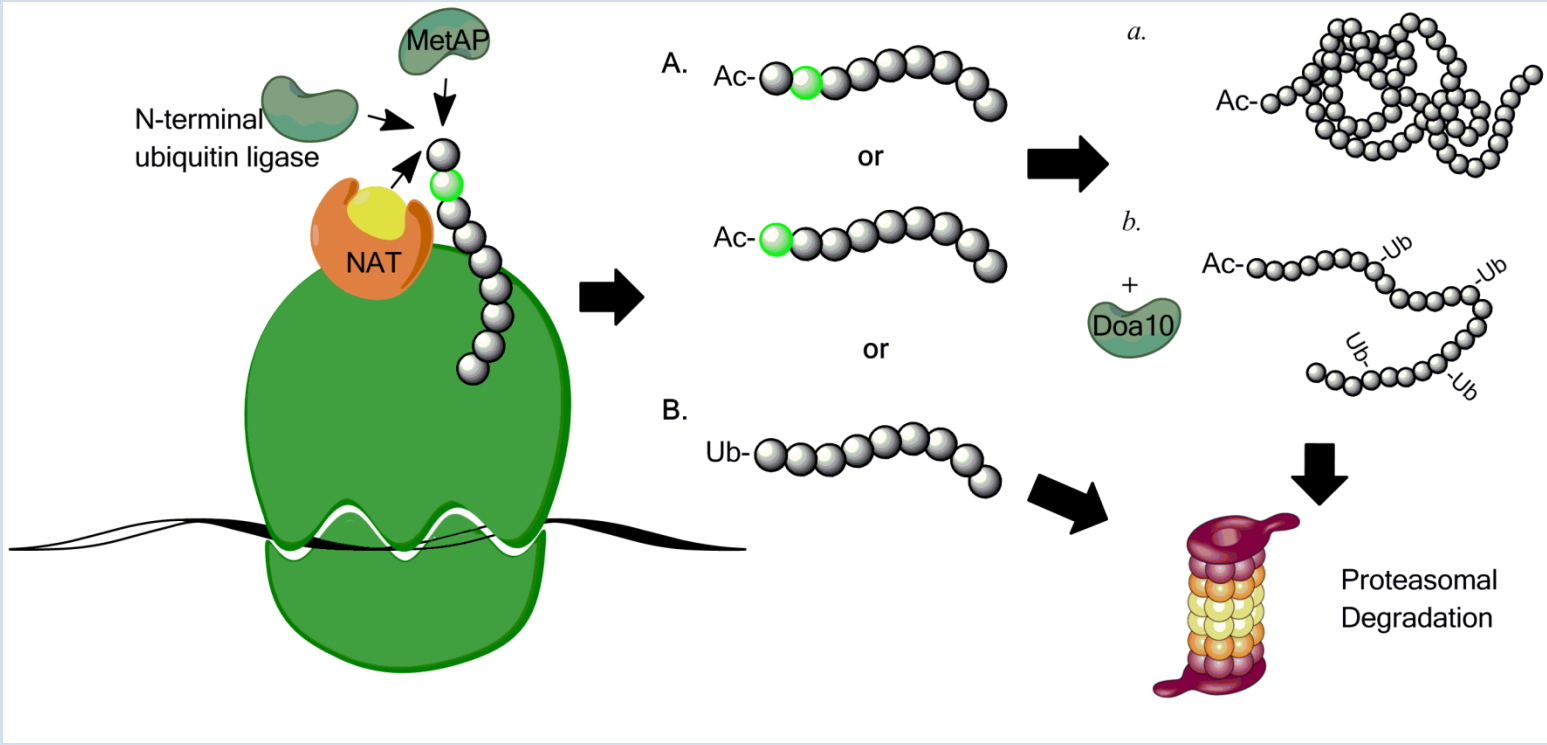


Figure courtesy of Kris Gevaert

N-Terminal Acetylation of Cellular Proteins Creates Specific Degradation Signals

Cheol-Sang Hwang, Anna Shemorry, Alexander Varshavsky*

N-Terminal Acetylation Inhibits Protein Targeting to the Endoplasmic Reticulum

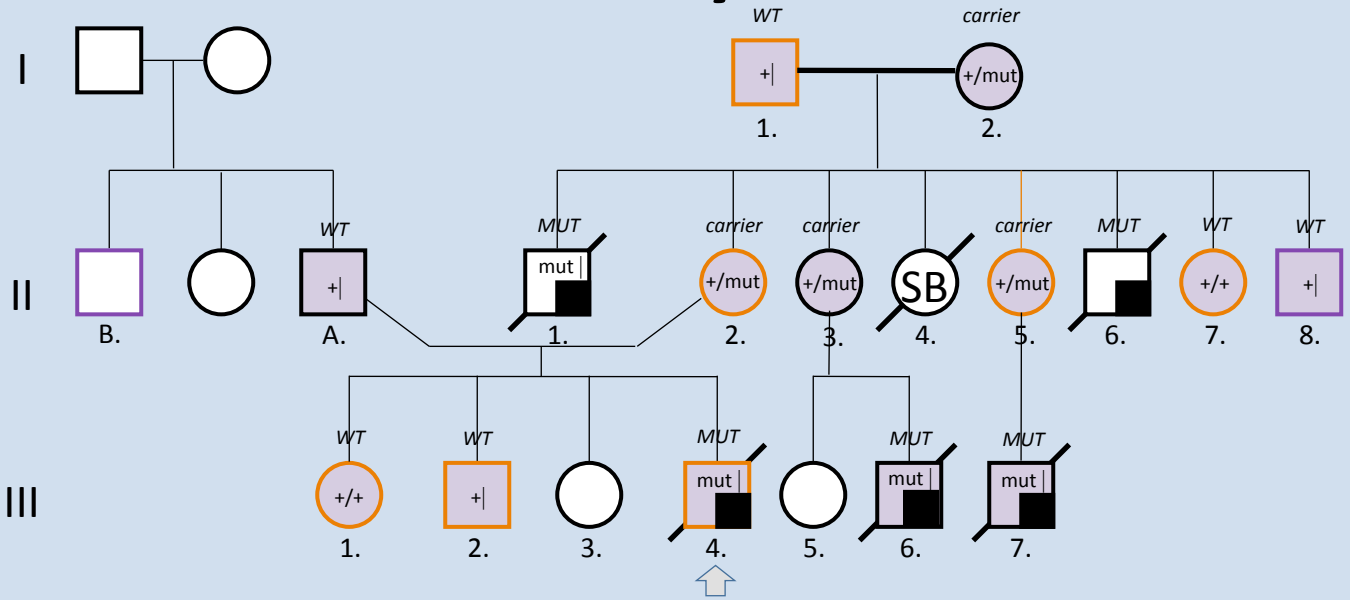
Gabriella M. A. Forte, Martin R. Pool*, Colin J. Stirling*

Faculty of Life Sciences, University of Manchester, Manchester, United Kingdom

Abstract

Amino-terminal acetylation is probably the most common protein modification in eukaryotes with as many as 50%–80% of proteins reportedly altered in this way. Here we report a systematic analysis of the predicted N-terminal processing of cytosolic proteins versus those destined to be sorted to the secretory pathway. While cytosolic proteins were profoundly biased in favour of processing, we found an equal and opposite bias against such modification for secretory proteins. Mutations in secretory signal sequences that led to their acetylation resulted in mis-sorting to the cytosol in a manner that was dependent upon the N-terminal processing machinery. Hence N-terminal acetylation represents an early determining step in the cellular sorting of nascent polypeptides that appears to be conserved across a wide range of species.

Proteomics Analysis of EBV-transformed cell lines 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)**

N-terminal COFRADIC

PROTOCOL

Selecting protein N-terminal peptides by combined fractional diagonal chromatography

An Staes[1,2], Francis Impens[1,2], Petra Van Damme[1,2], Bart Ruttens[1,2], Marc Goethals[1,2], Hans Demol[1,2], Evy Timmerman[1,2], Joël Vandekerckhove[1,2] & Kris Gevaert[1,2]

[1]Department of Medical Protein Research, Vlaams Instituut voor Biotechnologie (VIB), Ghent, Belgium. [2]Department of Biochemistry, Ghent University, Ghent, Belgium. Correspondence should be addressed to K.G. (kris.gevaert@vib-ugent.be).

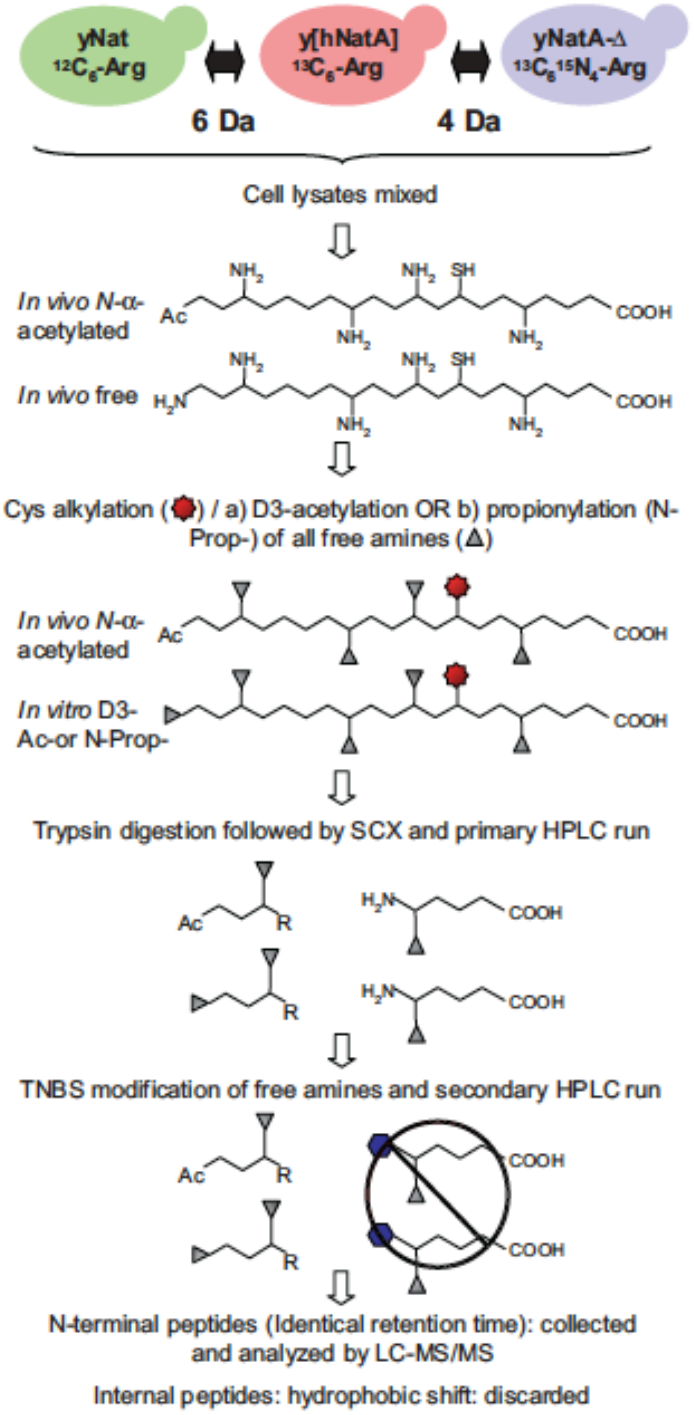
Published online 14 July 2011; doi:10.1038/nprot.2011.355

In recent years, procedures for selecting the N-terminal peptides of proteins with analysis by mass spectrometry have been established to characterize protease-mediated cleavage and protein α-*N*-acetylation on a proteomic level. As a pioneering technology, N-terminal combined fractional diagonal chromatography (COFRADIC) has been used in numerous studies in which these protein modifications were investigated. Derivatization of primary amines—which can include stable isotope labeling—occurs before trypsin digestion so that cleavage occurs after arginine residues. Strong cation exchange (SCX) chromatography results in the removal of most of the internal peptides. Diagonal, reversed-phase peptide chromatography, in which the two runs are separated by reaction with 2,4,6-trinitrobenzenesulfonic acid, results in the removal of the C-terminal peptides and remaining internal peptides and the fractionation of the sample. We describe here the fully matured N-terminal COFRADIC protocol as it is currently routinely used, including the most substantial improvements (including treatment with glutamine cyclotransferase and pyroglutamyl aminopeptidase to remove pyroglutamate before SCX, and a sample pooling scheme to reduce the overall number of liquid chromatography—tandem mass spectrometry analyses) that were made since its original publication. Completion of the N-terminal COFRADIC procedure takes ~5 d.

Staes A *et al.* (**2011**) *Nat. Protoc.* **6**, 1130-1141

Gevaert K *et al.* (**2003**) *Nat. Biotechnol.* **21**, 566-569

Proteomics Strategy With Thomas Arnesen, Petra van Damme And Kris Gevaert



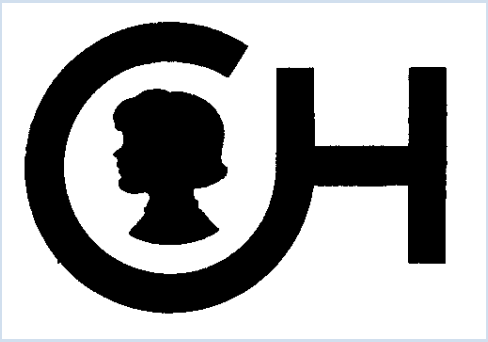
Results from EBV-transformed lymphocytes

AAAEEEDGGPEGPNR	66	86	92	91	87	84	91	Q99942	RNF5_HUMAN	E3 ubiquitin-protein ligase RNF5	Membrane; Multi-pass membrane protein. Mitochondrion membrane. Endoplasmic reticulum membrane.
AADTQVSETLKR	52	80	84	84	80	81	85	Q92616	GCN1L_HUMAN	Translational activator GCN1	
AAESALQVVEKLQAR	58	89	92	92	87	87	92	Q14241	ELOA1_HUMAN	Transcription elongation factor B polypeptide 3	Nucleus.
AVFADLDLR	66	95	96	96	96	95	100	P78346	RPP30_HUMAN	Ribonuclease P protein subunit p30	Nucleus;nucleolus.
MVEKEEAGGGISEEEAAQYDR	69	90	91	94	91	95	96	Q9UBE0	SAE1_HUMAN	SUMO-activating enzyme subunit 1	Nucleus.
MLGAPDESSVR	51	79	74	79	70	72	80	Q7Z4S6	KI21A_HUMAN	Kinesin-like protein KIF21A	Cytoplasm;cytoskeleton.
MLSPEAER	74	97	97	97	96	97	97	Q9NUG6	PDRG1_HUMAN	p53 and DNA damage-regulated protein 1	Cytoplasm.
AAGGGGGSSKASSSSASSAGALESSLDR	72	85	84	85	82	84	87	Q5VT52	RPRD2_HUMAN	Regulation of nuclear pre-mRNA domain-containing protein 2	
GEEANDDKKPTTKFELER	79	91	93	93	94	93	92	Q92989	CLP1_HUMAN	Polyribonucleotide 5'-hydroxyl-kinase Clp1	Nucleus.

My goal is to study N-terminal acetylation as follows:

- X-chromosome skewing assays on DNA from blood in the carrier females showed extreme skewing in all carriers. Therefore, lymphocytes without the mutation may have selective advantage? Why?
- Proteomics screens, identifying thus far ~40 proteins that are substantially hypo-acetylated in the proband's EBV-transformed lymphocytes vs. all other members of family.
- Skin fibroblasts from proband obtained, and now making iPS cells, enabling other studies. Access to other tissues possible.
- Knockout Mouse for *NAA10*, along with Knock-in for Naa10p-S37P, to study in this organism.
- Morpholino knockdown of NAA10 in zebrafish to recapitulate phenotype in zebrafish, allowing studies in this organism as well.

Utah Foundation for Biomedical Research



INFORMED CONSENT AUTHORIZATION TO PARTICIPATE IN A CLINICAL INVESTIGATION

Family Name:_____

Title: (Protocol #: 100) Study of the Genetic Causes of Complex Neurologic Psychiatric Disorders

Version: 14-Apr-2011
Protocol: 100

APPROVED BY
Independent IRB

Signature

14-Apr-2011
Date

This is just one example of many rare Mendelian disorders yet to be discovered

Long-range Plans: ~750 DNA samples from many pedigrees with 357 genotyped at CHOP thus far on Illumina 610K arrays

Table 1. Characteristics of seven new Utah extended pedigrees with preliminary diagnostic information.

Pedigree	# generations	# with DNA	# TS	# CMT	# CVT	# OCD*	# sub OCD**
14349	4	65	13	7	5	29	14
7166	3	27	7	1	0	11	10
13166	3	23	10	2	1	3	6
8115	3	20	9	1	0	9	3
6991	4	15	8	2	0	4	2
8598	3	11	8	0	0	6	0
3695	3	7	3	1	0	4	0
TOTALS		**168**	**58**	**14**	**6**	**66**	**35**

Note. TS=Tourette Syndrome; CMT=Chronic Motor Tics; CVT=Chronic Vocal Tics; OCD= Obsessive Compulsive Disorder; sub OCD=subclinical Obsessive Compulsive Disorder.

*Of the cases with OCD, 39 also have TS or chronic tics, leaving 27 with OCD only.

**Of the cases with sub OCD, 17 also have TS or chronic tics, leaving 18 with sub OCD only.

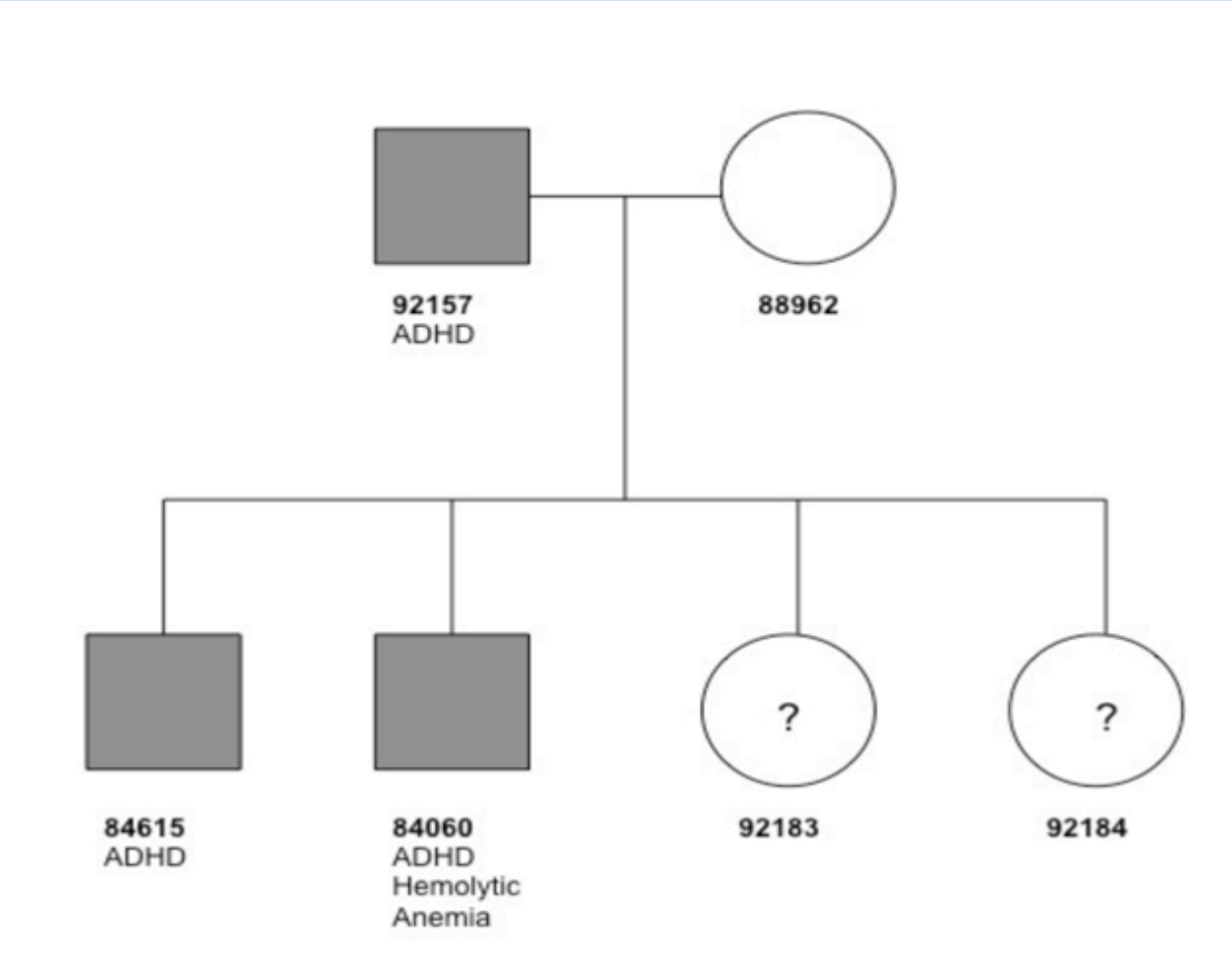
Genetics of Complex Neuropsychiatric Disorders is MUCH harder and will requires decades of more effort

Exome Sequencing and Unrelated Findings in the Context of Complex Disease Research: Ethical and Clinical Implications

GHOLSON J. LYON, TAO JIANG, RICHARD VAN WIJK, WEI WANG, PAUL MARK BODILY, JINCHUAN XING, LIFENG TIAN, REID J. ROBISON, MARK CLEMENT, LIN YANG, PENG ZHANG, YING LIU, BARRY MOORE, JOSEPH T. GLESSNER, JOSEPHINE ELIA, FRED REIMHERR, WOUTER W. VAN SOLINGE, MARK YANDELL, HAKON HAKONARSON, JUN WANG, WILLIAM EVAN JOHNSON, ZHI WEI, AND KAI WANG

Discov Med. 2011 Jul;12(62):41-55.

This is one ADHD pedigree in January 2010, very well phenotyped for four members



Discov Med. 2011 Jul;12(62):41-55.

Table 4: validated variants for ADHD and their population frequency in 5,680 and ~600 deep-sequenced exomes

# chrom	Position in HG19	Reference allele	Mutant allele	Gene	Type of Mutation	Amino acid change	# variants in 5680 BGI exomes[1]	% in BGI exomes	# variants in ~600 Baylor exomes	% in Baylor exomes
chr17	66872692	A	G	ABCA8	Nonsynonymous	C1387R	0	0.0%	0	0.0%
chr11	68566802	G	A	CPT1A	Nonsynonymous	L193F	0	0.0%	0	0.0%
chr8	100994274	A	G	RGS22	Nonsynonymous	I1084T	0	0.0%	0	0.0%
chr18	61654247	G	T	SERPINB8	Nonsynonymous	G287V	0	0.0%	0	0.0%
chr1	207200877	-	T	C1orf116	frameshift insertion		34	1.4%	0	0.0%
chr18	29101156	T	G	DSG2	Nonsynonymous	V158G	1	0.0%	1	0.2%
chr3	125877290	**G**	**A**	**ALDH1L1**	**Nonsynonymous**	**P107L**	**2**	**0.0%**	**0**	**0.0%**
chr13	52542680	**A**	**G**	**ATP7B**	**Nonsynonymous**	**V536A**	**1**	**0.0%**	**1**	**0.2%**
chr10	53458646	**A**	**C**	**CSTF2T**	**Nonsynonymous**	**C222G**	**4**	**0.1%**	**1**	**0.2%**
chr14	21972019	**G**	**A**	**METTL3**	**Nonsynonymous**	**R36W**	**9**	**0.2%**	**1**	**0.2%**
chr11	76954790	-	A	GDPD4	frameshift insertion		36	1.5%	6	1.0%
chr7	87160618	A	T	ABCB1	Nonsynonymous	S893T	815	14.3%[1]	9	1.5%
chr11	134128923	C	G	ACAD8	Nonsynonymous	S171C	112	2.0%	20	3.3%
chr20	17956347	C	T	C20orf72	Nonsynonymous	R178W	23	0.4%	8	1.3%
chr8	33318891	T	C	FUT10	Nonsynonymous	Q27R	15	0.3%	3	0.5%
chr13	20797025	A	T	GJB6	Nonsynonymous	S199T	68	1.2%	4	0.7%
chr16	71015329	G	T	HYDIN	Nonsynonymous	P1491H	77	1.4%	dozens	>5.0%
chr10	22019855	G	A	MLLT10	Nonsynonymous	R713H	15	0.3%	6	1.0%
chr17	10415269	A	G	MYH1	Nonsynonymous	Y435H	99	1.7%	14	2.3%
chr1	145015877	G	T	PDE4DIP	Nonsynonymous	L142I	1256	22.1%	hundreds	>30.0%
chr2	98809432	T	C	VWA3B	Nonsynonymous	I513T	15	0.3%	16	2.7%
chr5	115202418	AAGA	-	AP3S1	frameshift deletion		185	7.8%	19	3.2%

1. The indels were only measured thus far in 2360 exomes at BGI, whereas the SNPs were measured in 5680 exomes.

Exome Sequencing performed early 2010

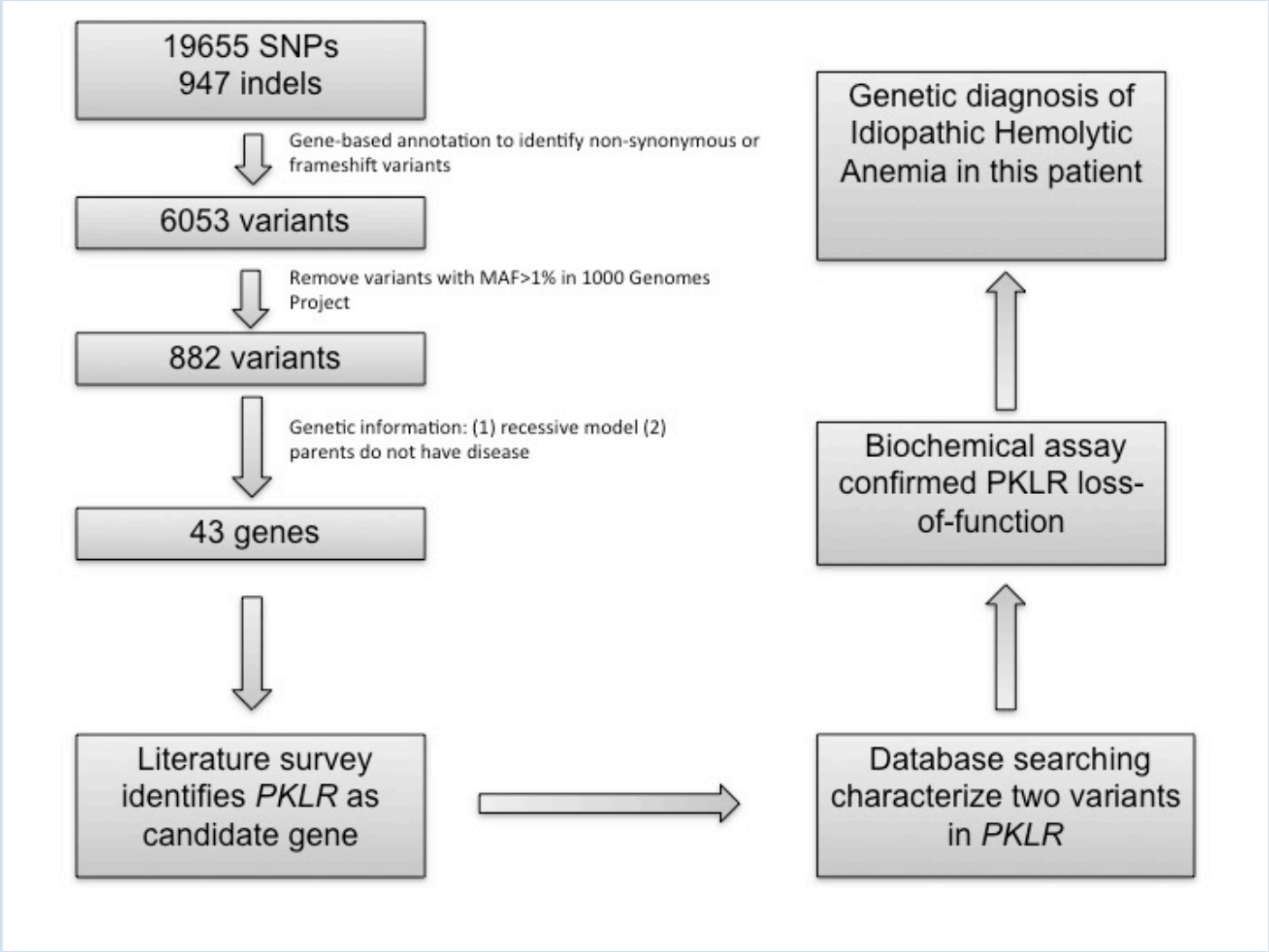
While analyzing the exome data, research subject (age ~24) informs me that he recently had his spleen removed!

He has idiopathic hemolytic anemia, since childhood….

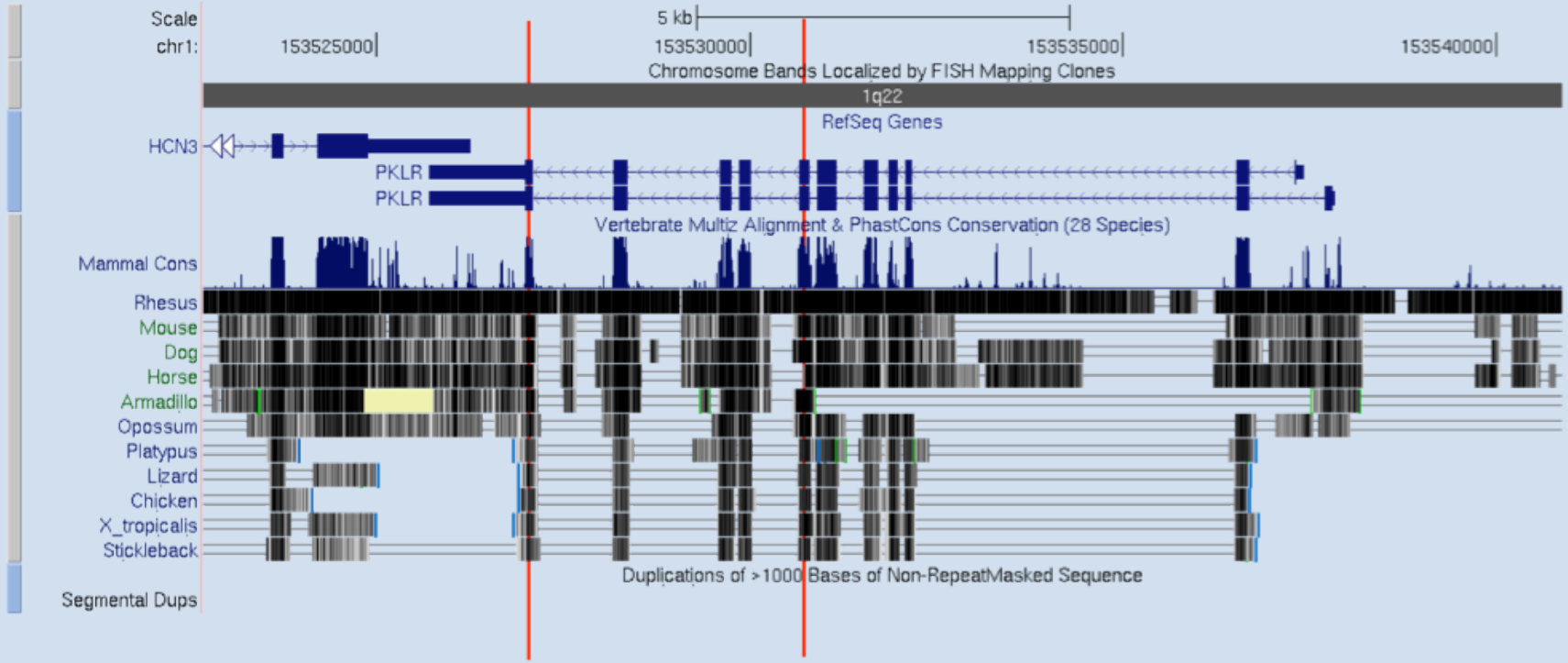
Although I am not his physician, I still feel an ethical and moral obligation to try to figure out what is going on….

A corollary of this is that I am actively working on the ethical and clinical aspects of sequencing human genomes.

This was our filtering pipeline



Compound Heterozygote in *PKLR*, with each mutation inherited from one parent



Some Additional Data to support the causation of these variants for idiopathic hemolytic anemia

Table 2: Biochemical assays of enzyme activities in the patient affected with idiopathic hemolytic anemia confirmed *PKLR* deficiency. PK, pyruvate kinase; HK, hexokinase; G6PD, glucose-6-phosphate dehydrogenase.

	Patient 84060	Control	Reference values
PK (U/gHb)	3.3 L	8.6	6.1 – 12.3
HK (U/gHb)	3.2 H	1.1	0.8 – 1.5
G6PD (U/gHb)	15.8 H	9.2	6.4 – 10.5

Table 3: Bioinformatics prediction on the functional impact of two PKLR mutations. A mutation is regarded as deleterious if the SIFT<0.05, or PolyPhen>0.85, or PhyloP>0.95, or MutationTaster/LRT prediction as "D" (deleterious).

Mutation	SIFT	PolyPhen 2	PhyloP	LRT	MutationTaster
R569Q	0.03	0.84	0.97	D	D
G341A	0	0.889	1	D	D

Structural Modeling is also consistent with deleterious effects of these mutations.

Proving CAUSALITY for variants is critically important

- We amassed evidence for causality of the PKLR mutations.
- We found many very rare variants segregating with ADHD in father and two sons, which we replicated with second round of exome capture and sequencing followed by VAAST analysis on same samples.
- BUT we have not yet proven with certainty that any of these variants cause the disease, either by themselves, or in aggregate, or with epistatic effects.
- Biology experiments and case-control genotyping must be performed to further bolster causality.

And there are so many more intriguing psychiatric diseases to discover and work on!

Video…..

Case Presentation

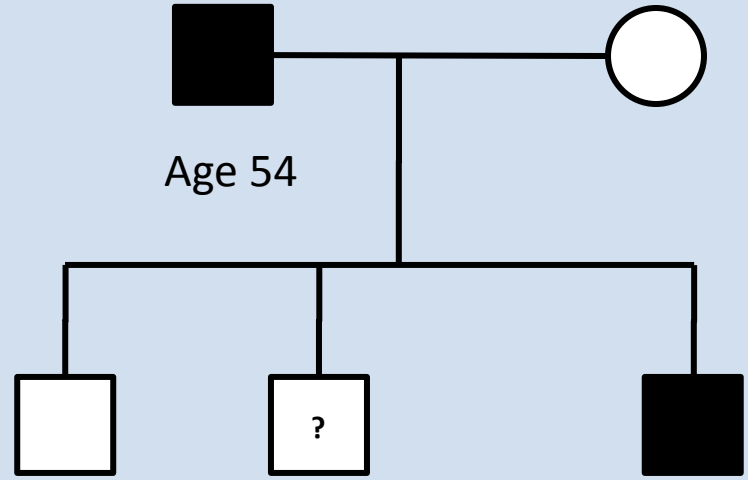
- Male, age 55 currently.
- Psychotic break at age 20 with bipolar features.
- Evolution into schizoaffective disorder over next 25 years.
- Also with severe obsessive compulsive disorder and severe Tourette Syndrome
- At least two very severe suicide attempts at age 22, including throwing self under a truck one time and then driving head-on into another car (with death of two passengers in other car, found not guilty by reason of insanity).

- Extensive medication trials over many years, along with anterior capsulotomy with very little effect for the OCD.

- Current meds:

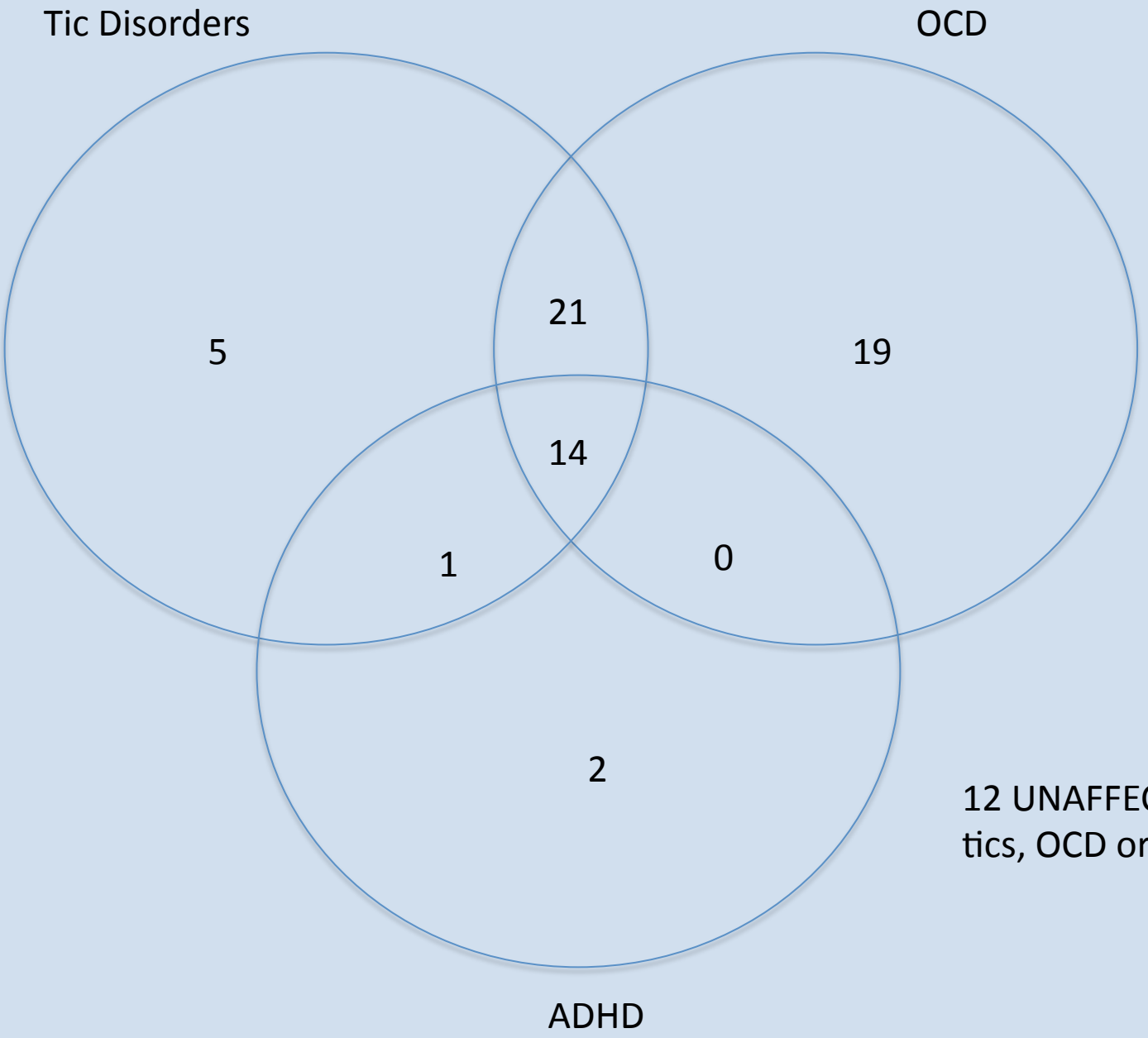
Klonopin	Lithium
Nicotinamide	Seroquel
Lunesta	Lamictal
Ativan	Luvox

Expanding the Pedigree –K8101



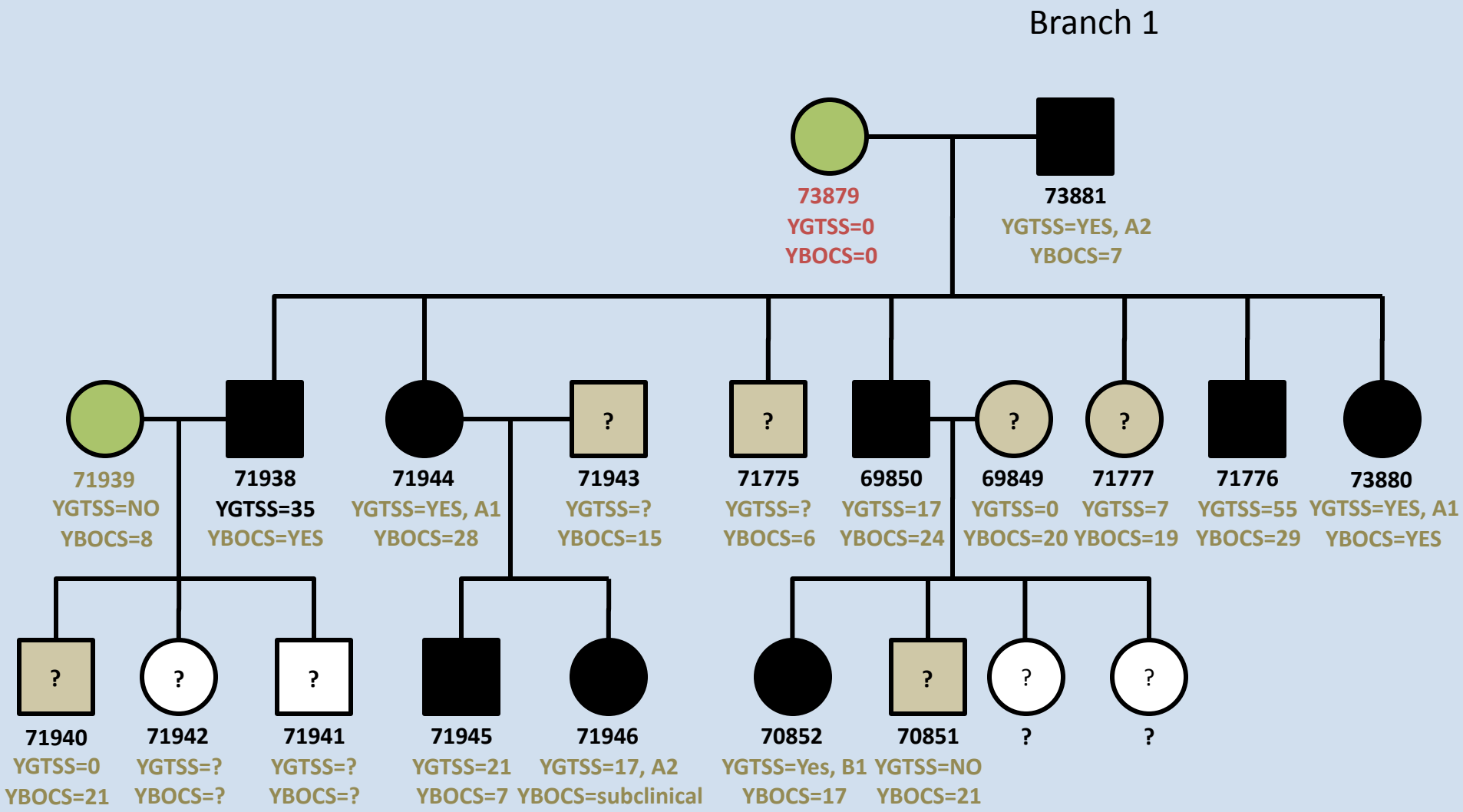
Collected 35 DNA samples from the extended family, due to very large excess of major depression, bipolar, Tourette and OCD.

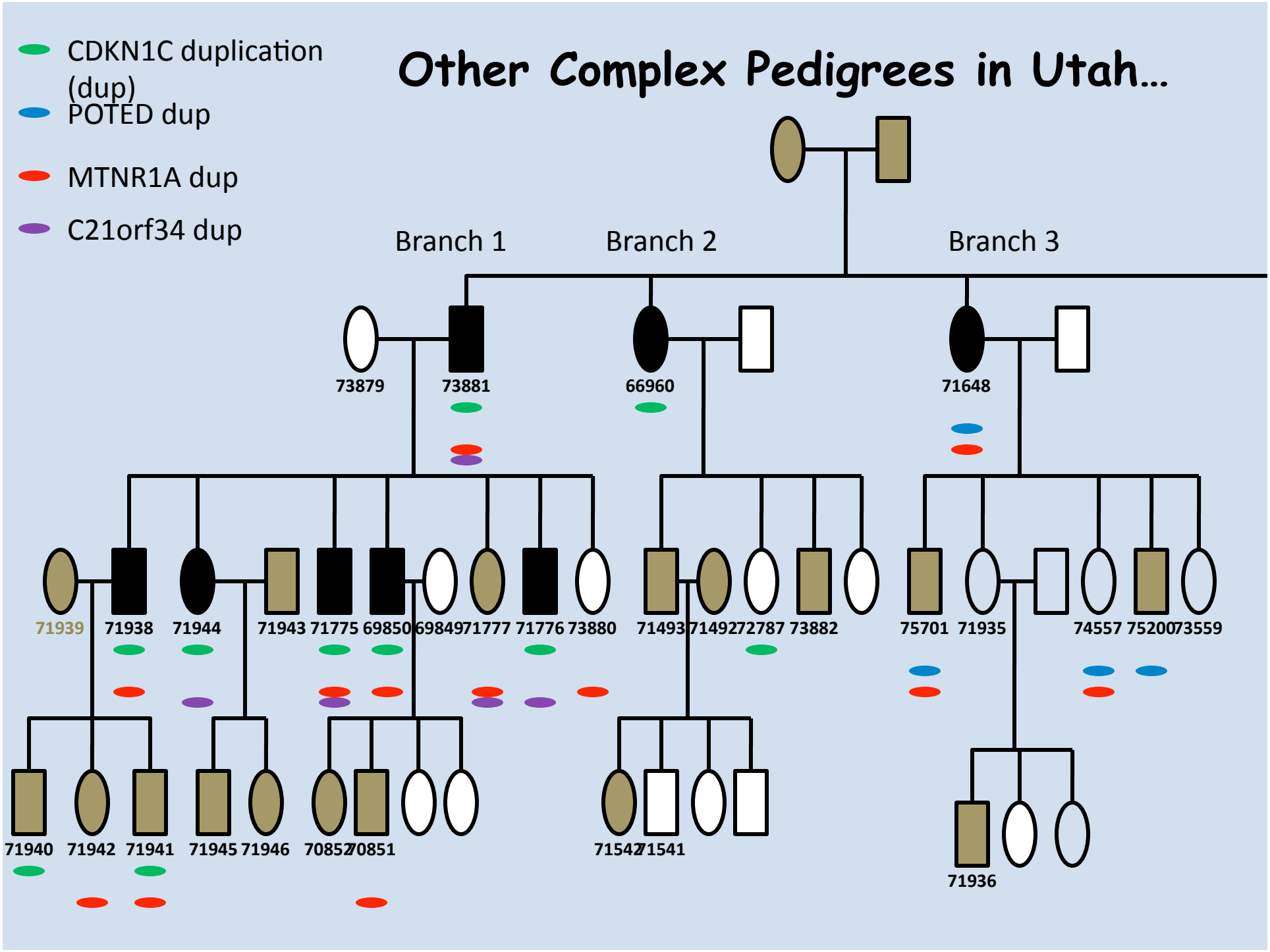
Another Large Pedigree in Utah with 74 members

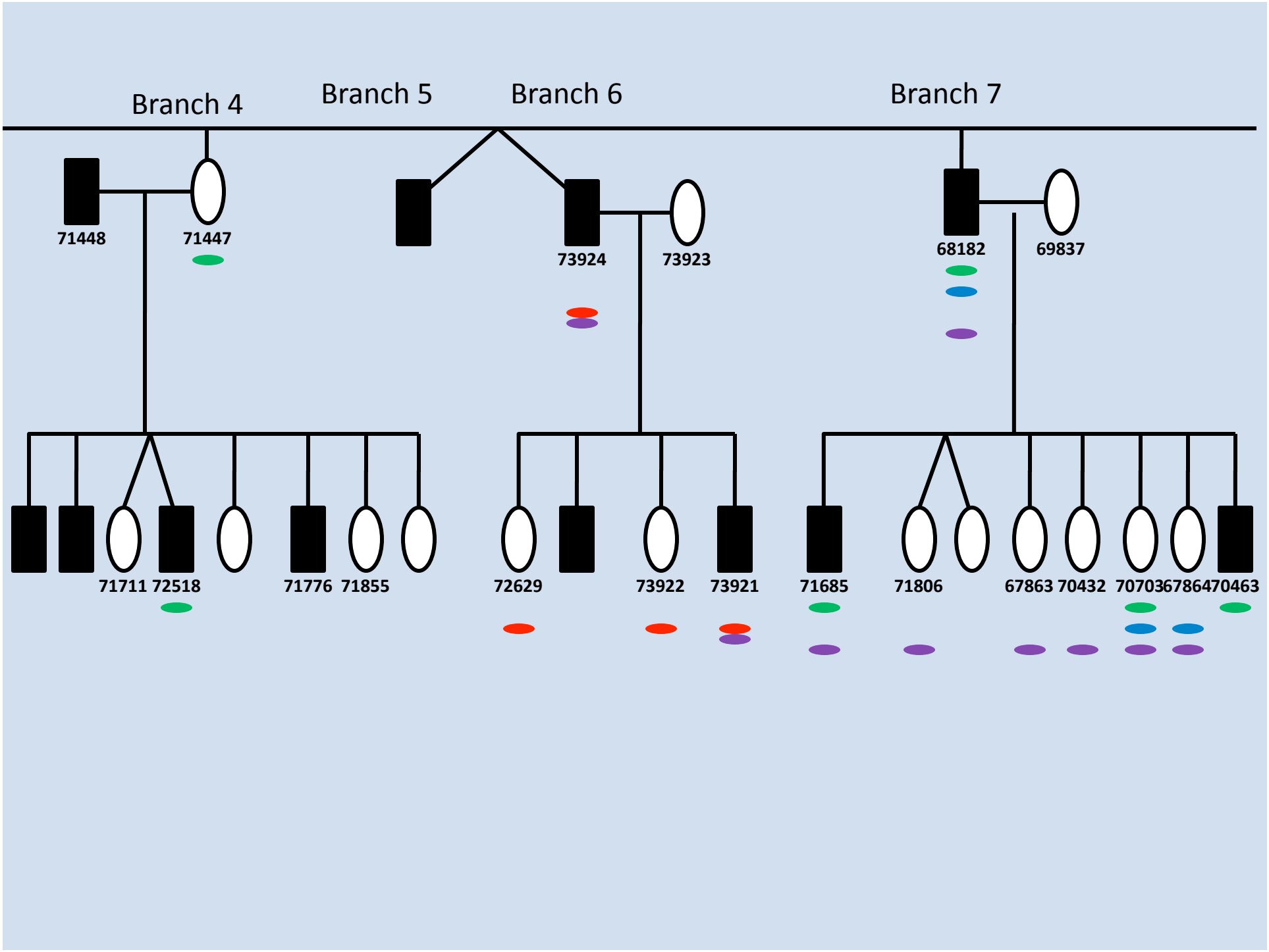


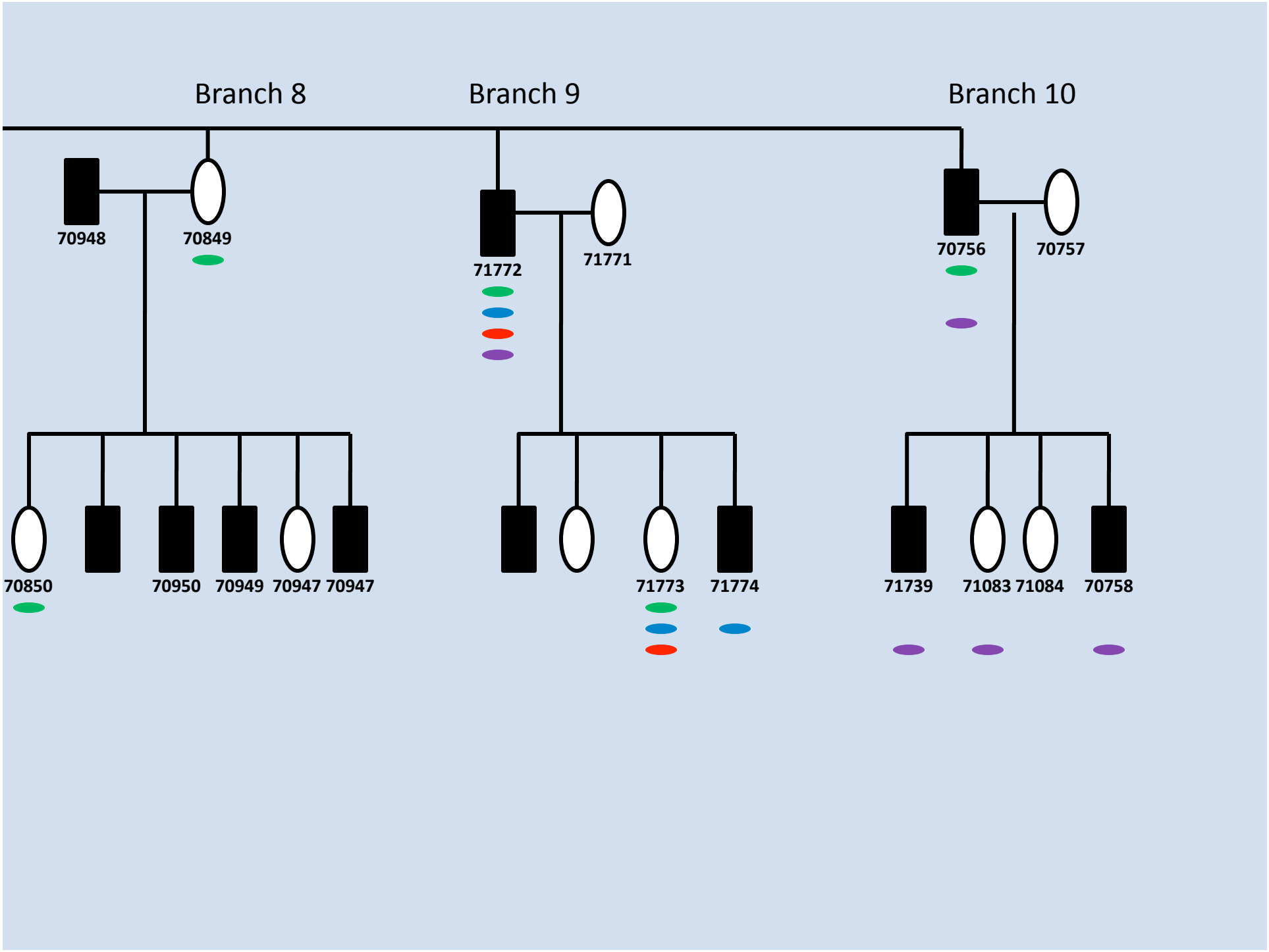
12 UNAFFECTED with tics, OCD or ADHD

Phenotyping of just one branch in this pedigree




CDKN1C duplication (dup)
POTED dup
MTNR1A dup
C21orf34 dup
Other Complex Pedigrees in Utah…
Branch 1
Branch 2
Branch 3
73879
73881
66960
71648
71939
71938
71944
71943
71775
69850
69849
71777
71776
73880
71493
71492
72787
73882
75701
71935
74557
75200
73559
71940
71942
71941
71945
71946
70852
70851
71542
71541
71936


Branch 4
Branch 5
Branch 6
Branch 7
71448
71447
73924
73923
68182
69837
71711
72518
71776
71855
72629
73922
73921
71685
71806
67863
70432
707036
786
470463


Branch 8
Branch 9
Branch 10
70948
70849
71772
71771
70756
70757
70850
70950
70949
70947
70947
71773
71774
71739
71083
71084
70758

We have now sequenced 15 exomes with 44MB Agilent capture kit and sequenced to attain goal of >80% of target region at >20 reads per bp

Exome Capture Statistics	K24510-84060	K24510-92157-a	K24510-84615	K24510-88962
Target region (bp)	46,401,121	46,401,121	46,401,121	46,257,379
Raw reads	138,779,950	161,898,170	156,985,870	104,423,704
Raw data yield (Mb)	12,490	14,571	14,129	9,398
Reads mapped to genome	110,160,277	135,603,094	135,087,576	83,942,646
Reads mapped to target region	68,042,793	84,379,239	80,347,146	61,207,116
Data mapped to target region (Mb)	5,337.69	6,647.18	6,280.01	4,614.47
Mean depth of target region	**115.03**	**143.25**	**135.34**	**99.76**
Coverage of target region (%)	**0.9948**	**0.9947**	**0.9954**	**0.9828**
Average read length (bp)	89.91	89.92	89.95	89.75
Fraction of target covered >=4X	98.17	98.38	98.47	94.25
Fraction of target covered >=10X	95.18	95.90	95.97	87.90
Fraction of target covered >=20X	**90.12**	**91.62**	**91.75**	**80.70**
Fraction of target covered >=30X	84.98	87.42	87.67	74.69
Capture specificity (%)	61.52	62.12	59.25	73.16
Fraction of unique mapped bases on or near target	65.59	65.98	63.69	85.46
Gender test result	M	M	M	F

The current state of Psychiatric Genetics

Molecular Psychiatry (2011), 1–12
© 2011 Macmillan Publishers Limited All rights reserved 1359-4184/11
www.nature.com/mp

EXPERT REVIEW

Evidence-based psychiatric genetics, AKA the false dichotomy between common and rare variant hypotheses

PM Visscher[1], ME Goddard[2,3], EM Derks[4] and NR Wray[1]

[1]*Queensland Institute of Medical Research, Herston, Queensland, Australia;* [2]*Department of Primary Industries, Biosciences Research Division, Melbourne, Victoria, Australia;* [3]*Department of Agriculture and Food Systems, University of Melbourne, Melbourne, Australia and* [4]*Neuroimaging Research Group, University Medical Center, Utrecht, The Netherlands*

"From the empirical evidence on segregation patterns within families, the recurrence risk to relatives and the evidence from GWAS and CNV analyses, it appears beyond doubt that most psychiatric disease is not Mendelian in the sense that they are not caused solely by a single mutation (others disagree with this conclusion[28,62]). That is not to say that Mendelian forms cannot exist, but if they do, they account for very little of the population variance."

Psychological Medicine (2011), **41**, 19–32. © Cambridge University Press 2010
doi:10.1017/S003329171000070X

REVIEW ARTICLE

Rethinking the genetic architecture of schizophrenia

K. J. Mitchell[1]* and D. J. Porteous[2]

[1] *Smurfit Institute of Genetics, Trinity College Dublin, Ireland*
[2] *Medical Genetics Section, University of Edinburgh Molecular Medicine Centre, Institute of Genetics and Molecular Medicine, Edinburgh, UK*

BRITISH JOURNAL OF PSYCHIATRY **(2007), 190, 194–199. doi: 10.1192/bjp.bp.106.025585**

Schizophrenia: a common disease caused by multiple rare alleles†

JON M. McCLELLAN, EZRA SUSSER and MARY-CLAIRE KING

Molecular Psychiatry (2009) 14, 740–745
© 2009 Nature Publishing Group All rights reserved 1359-4184/09 $32.00
www.nature.com/mp

LETTERS TO THE EDITOR

GWAS for psychiatric disease: is the framework built on a solid foundation?

Mitchell and Porteous

"A primary justification for any genetic study of a condition of uncertain etiology must surely be to shed light on the biological causes…. Focused studies on individual cases, single families or genetically homogeneous populations do not currently attract the same cash or cachet as consortium-based GWAS studies, but promise greater returns in terms of biological insight and etiological understanding.

….a pathway biology approach … provides a logical antidote to the uncertainty of ever larger, more heterogenous and more costly GWAS."

Future Directions

- Studies are ongoing with cell lines derived from the Ogden Syndrome family, including skin fibroblasts and EBV-transformed cell lines from blood-derived monocytes.
- Continue to collect DNA samples from families in Utah and throughout America through the Utah Foundation for Biomedical Research.
- Pursue the genetic basis of some rare disorders and also much more complex neuropsychiatric disorders, including Tourette Syndrome, autism, OCD, ADHD, and schizophrenia.
- Pursue the biology of new variants discovered.
- Collect and test DNA from more fetal and infantile deaths, as there will be much genetic variation implicated.

Final Thoughts

J. Rogers Hollingsworth: "The optimal environment for great innovation and scientific break-through is characterized by

1. a maximum of flexibility without hindrance by hierarchical structures,
2. a maximum of independence of researchers and strong encouragement for risky projects, and
3. a large variety of different cultural backgrounds of the researchers."

J.R. Hollingsworth. Institutionalizing Excellence in Biomedical Research. In: D.H. Stapelton, ed. *Creating a Tradition of Biomedical Research*, pp.17-63, New York 2004.

Acknowledgments

Alan Rope
John M. Opitz
John C. Carey
Steven Chin
Brian Dalley
Heidi Deborah Fain
Chad D. Huff
W. Evan Johnson
Lynn B. Jorde
Barry Moore
Theodore J. Pysher
Reid Robison
Christa Schank
Sarah T. South
Jeffrey J Swensen
Jinchuan Xing
Mark Yandell

Reid Robison
Edwin Nyambi
Matt Adams
Josh Hansen
Leone Lobendahn

Kai Wang
Hakon Hakonarson
Zhi Wei
Lifeng Tian
Josephine Elia

Evan Johnson
Paul Mark Bodily
Mark Clement

Leslie G. Biesecker
Jenifer J Johnston
Lynne M. Bird

Thomas Arnesen
Rune Evjenth
Johan R. Lillehaug

Richard Van Wijk
Wouter W. van Solinge

Tao Jiang
Jun Wang

our study families

The VAAST DEVELOPMENT GROUP

www.yandell-lab.org

- Mark Yandell
- Martin Reese
- Hao Hu[+]
- Barry Moore[+]
- Steve Chervitz[+!]
- Chad Huff[x]
- Jinchuan Xing[x+]
- Marc Singleton[+]
- Edward Kiruluta[!]
- Archie Russell[!]
- Fidel Salas[!]
- Ginger Guozhen Fan[+]

[+]Yandell lab, [!]Omicia, [x]Jorde Lab

Omicia

THE UNIVERSITY OF UTAH

Final Thoughts

J. Rogers Hollingsworth: "The optimal environment for great innovation and scientific break-through is characterized by

1. a maximum of flexibility without hindrance by hierarchical structures,
2. a maximum of independence of researchers and strong encouragement for risky projects, and
3. a large variety of different cultural backgrounds of the researchers."

J.R. Hollingsworth. Institutionalizing Excellence in Biomedical Research. In: D.H. Stapelton, ed. *Creating a Tradition of Biomedical Research*, pp.17-63, New York 2004.