CRISPR and Diagnostics:
Challenges and Strategies for Understanding Results from Sequencing including Variants of Unknown Significance

Josh Bonkowsky, MD, PhD
Department of Pediatrics
University of Utah School of Medicine

UNIVERSITY OF UTAH
HEALTH SCIENCES

Outline

1. Pediatric Neurology and the Diagnosis Problem
2. Diagnosis: Costs and NGS (Next-Generation Sequencing)
    1. Leukodystrophy as an example
3. Crispy Zebrafish (… CRISPR and Zebrafish)
4. Perils and Successes with CRISPR Modeling
    1. Neuromuscular Disease
    2. The nav1 problem
    3. Lou Gehrig's disease

Pediatric Neurology and the Diagnosis Problem

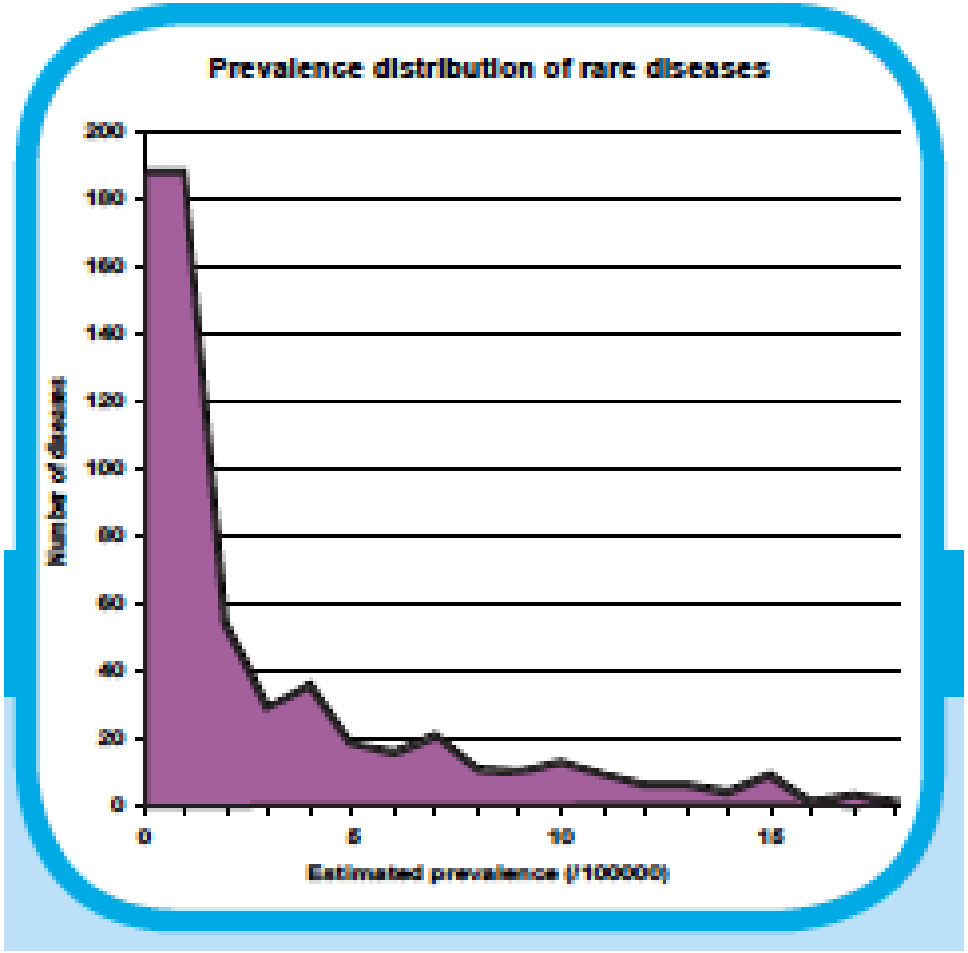
Pediatric Neurological Diseases

- ~5% of all children
- Life-long morbidity; higher mortality
- Largest single group of healthcare costs for children
  - contribution to the "Diagnostic Odyssey"
    - Berry, Poduri, Bonkowsky et al., 2012, *PLoS Medicine*
- Known and unknown causes of disease
  - many rare diseases
  - for most patients the genetic cause has been unknown

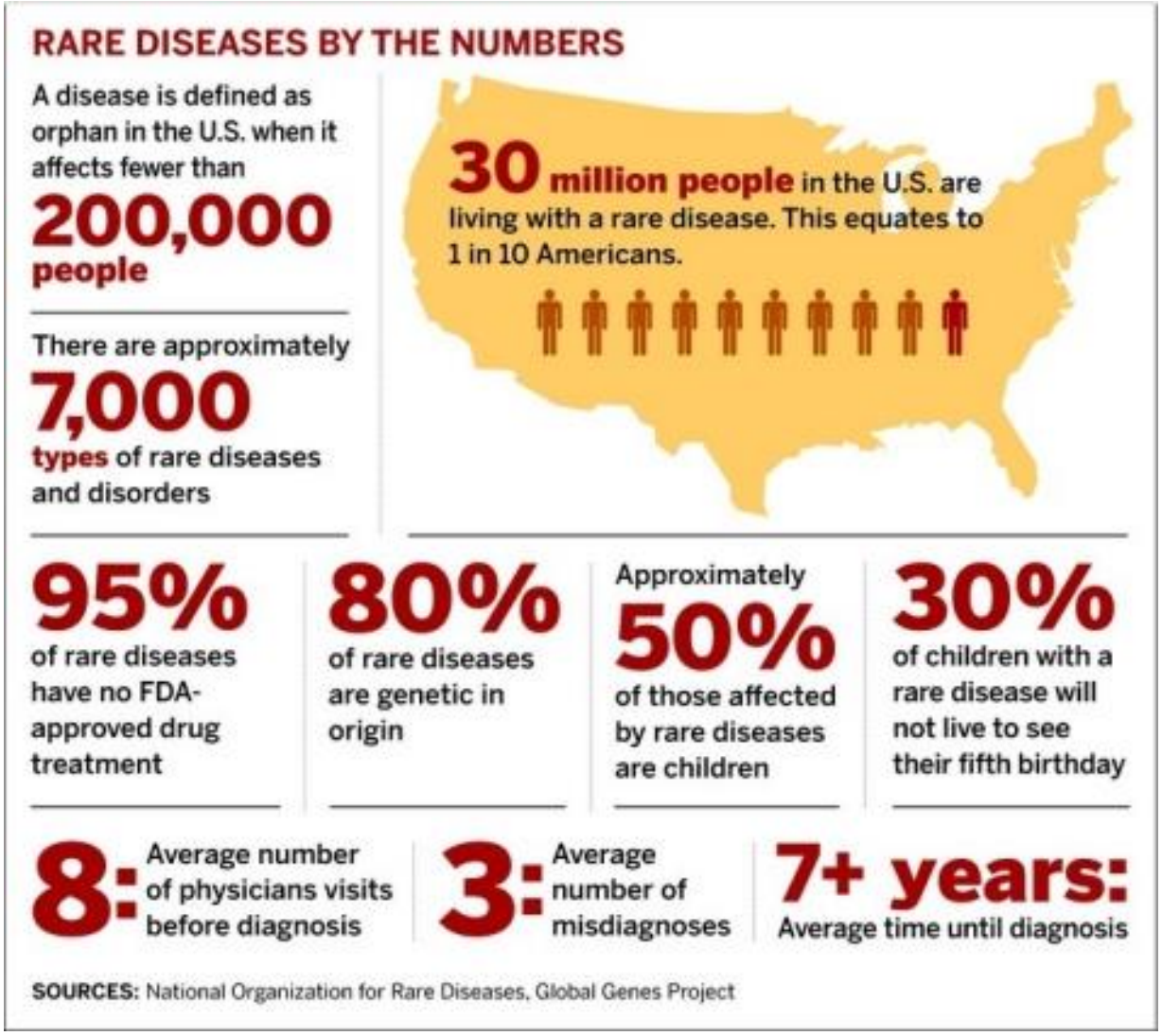
Rare and Orphan Diseases

- >2,025 rare diseases
- 25 million Americans affected
- orphan disease:

  "for which there is no reasonable expectation that the cost of developing and making available in the United States a drug for such disease or condition will [be] recovered from sales in the United States of such drug"



Orphan Diseases and Leukodystrophies



What is an undiagnosed disease?

- A disease that has not been diagnosed because the correct test has not yet been performed
  - rare disease
  - atypical presentation of a more common disease
- A disease that has not been diagnosed because we didn't know the disease existed
  - majority of undiagnosed diseases are neurologic

Why does diagnosis matter?

- Cure
- Therapy/Treatment
- Clinical Trials
- Natural history studies
- Prognosis for family
- Genetic counseling
- Genetic and biochemical pathways of disease

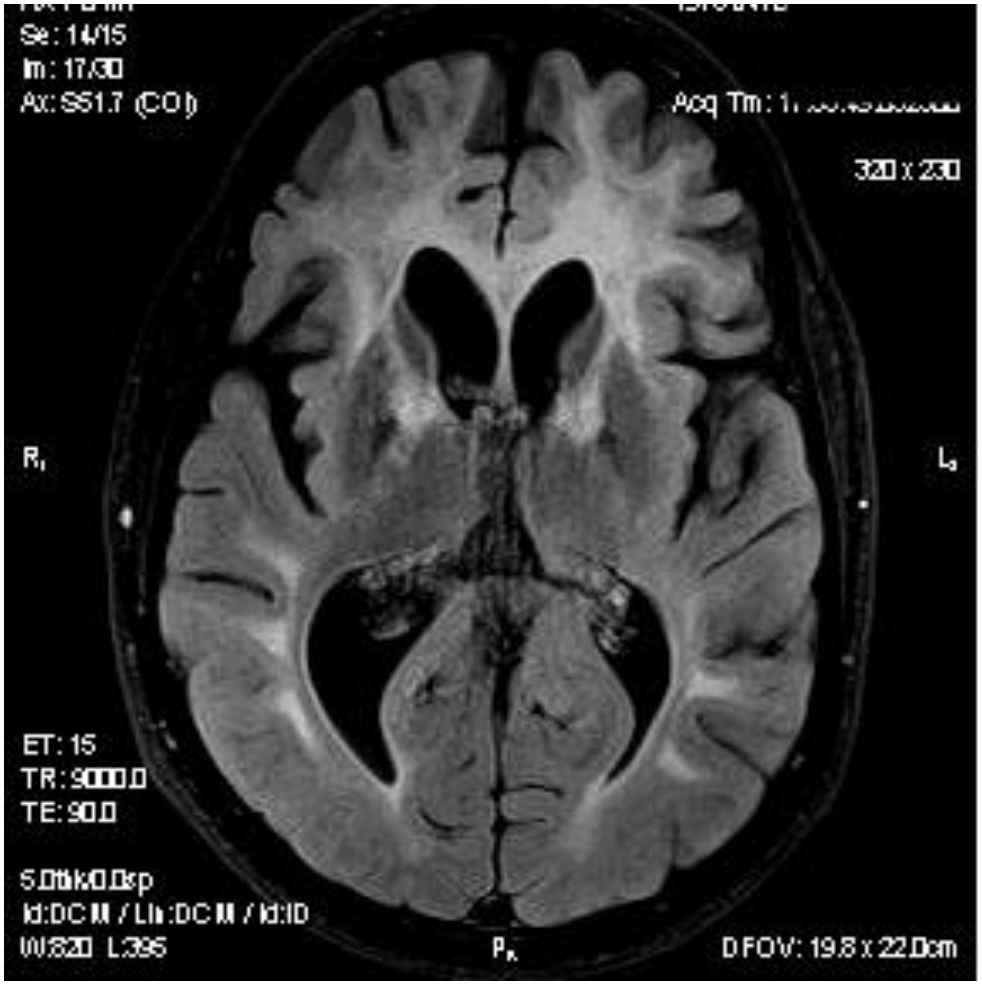
How good are we at diagnosis?

Pediatric Neurology

- MRI: 20% diagnosis
- CGH microarray: 10%
- NGS (Next-Generation Sequencing): 40%

Leukodystrophy

- Leukodystrophy:
  - Genetic
  - Involvement of white matter (myelin)
    - Not secondary to a different etiology (trauma, prematurity, etc.)

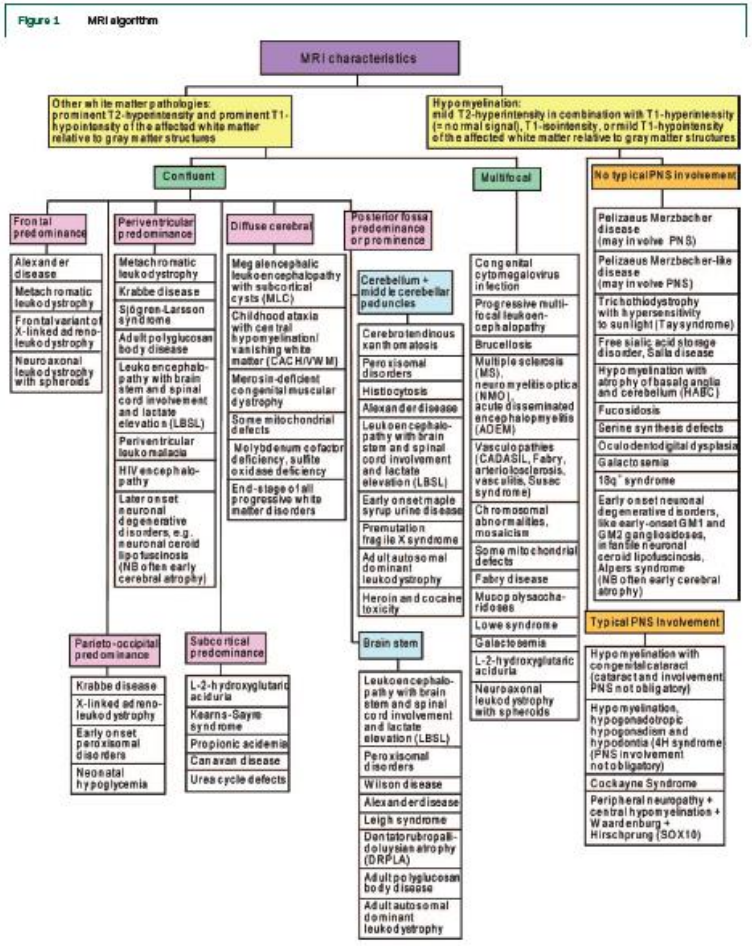


What is a Leukodystrophy?

- Three types:
  - Hypomyelination
  - Dysmyelination
  - Demyelination
- 30 canonical genes, >700 total genes
- Diagnosis rates ~50%

Leukodystrophy Problems

- Causes of leukodystrophies not known
- How to diagnose unknown
- No treatments

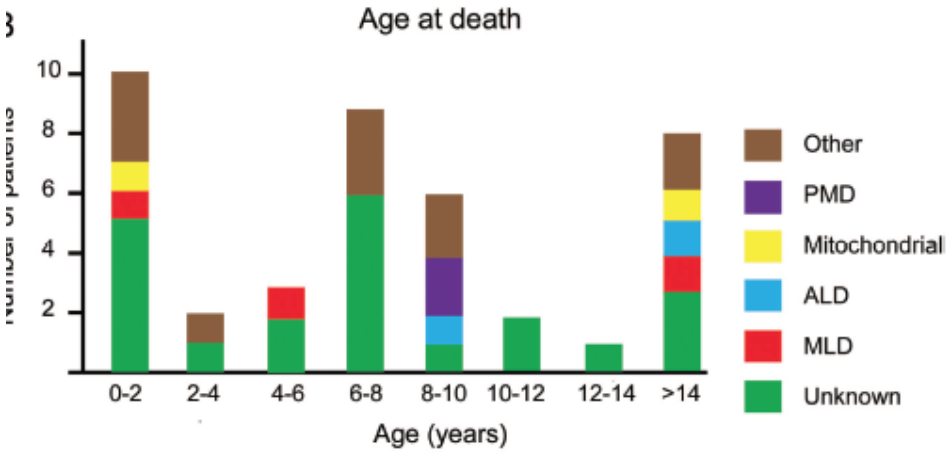



WESTERN
LEUKODYSTROPHY
PROJECT

The Burden of Leukodystrophies

Table 2 Death, neurologic features, and costs in the leukodystrophy cohort[a]

Outcomes	Values
Death, n (%)	42 (34)
Average age at death, y	8.2
Epilepsy, n (%)	60 (49)
Average age at onset, y	4.0
Developmental regression, n (%)	39 (32)
Feeding tube, n (%)	53 (43)
Costs	
Total cohort cost	$14,315,919
Average yearly cost/patient	$22,579



Bonkowsky et al., *Neurology*, 2010

Diagnosis: Costs and NGS (Next-Generation Sequencing)

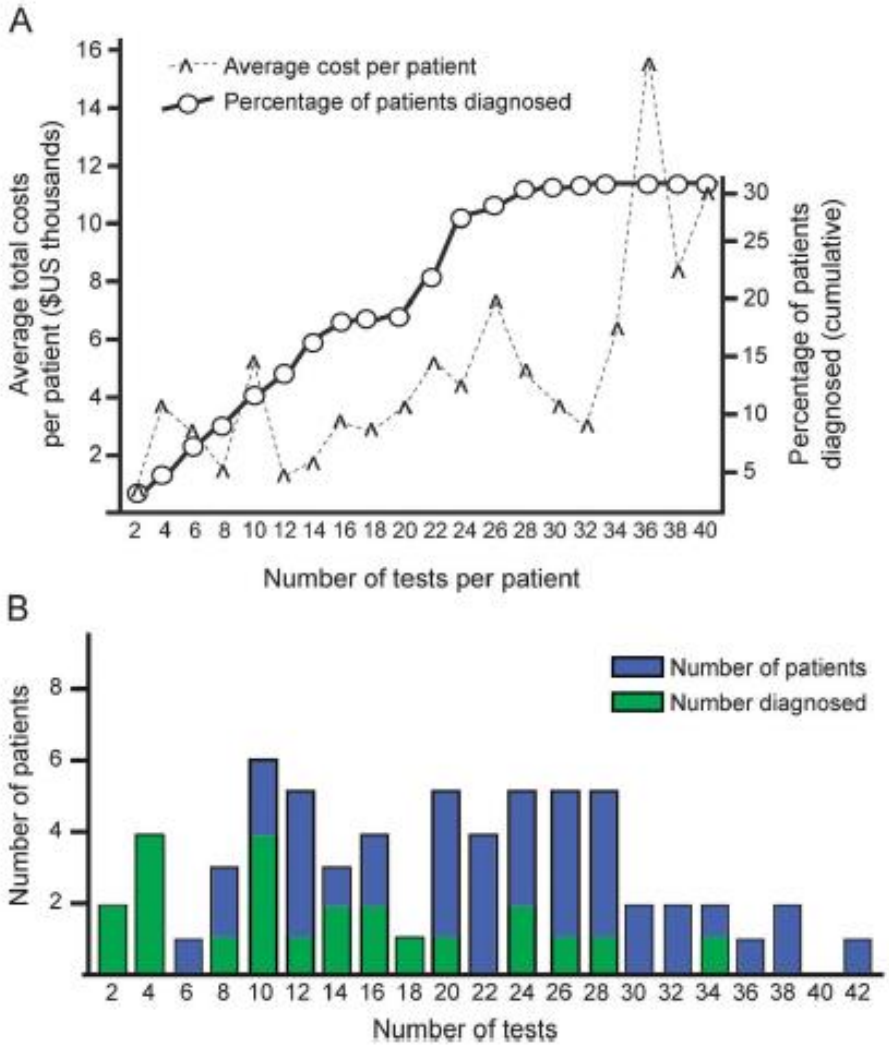
The Diagnostic Odyssey

Hypotheses:

1. costs are substantial.
2. NGS will help.

1. Costs are substantial

- **False**
- Average costs of $4209/patient
  - Compared to average healthcare costs of $107,000/patient
- Conclusion: *reaching a diagnosis is not the primary driver of costs*



2. NGS will help

- **True**
- Charges for the entire cohort= $538,053
- If NGS had been performed instead= $371,200
  - and equal or better diagnosis rate
- Conclusion: *Use NGS early*

Richards et al., 2015, Neurology
Richards et al., 2015, Am J Med Genetics

Next Generation Sequencing: NGS

- NGS has revolutionized diagnosis
  - Sequencing technology is on the time-scale of hours/days
    - Interpretation is weeks to months
- But accompanying limitations:
  - sequencing informatics bottleneck
  - ***biology bottleneck of variants***
    - each individual has ~74 germline de novo mutations
  - the spectre of non-coding variants
  - the role of somatic mutations

Diagnosis: Today!

Two Steps:

1. **Test treatable disorders**

Either:

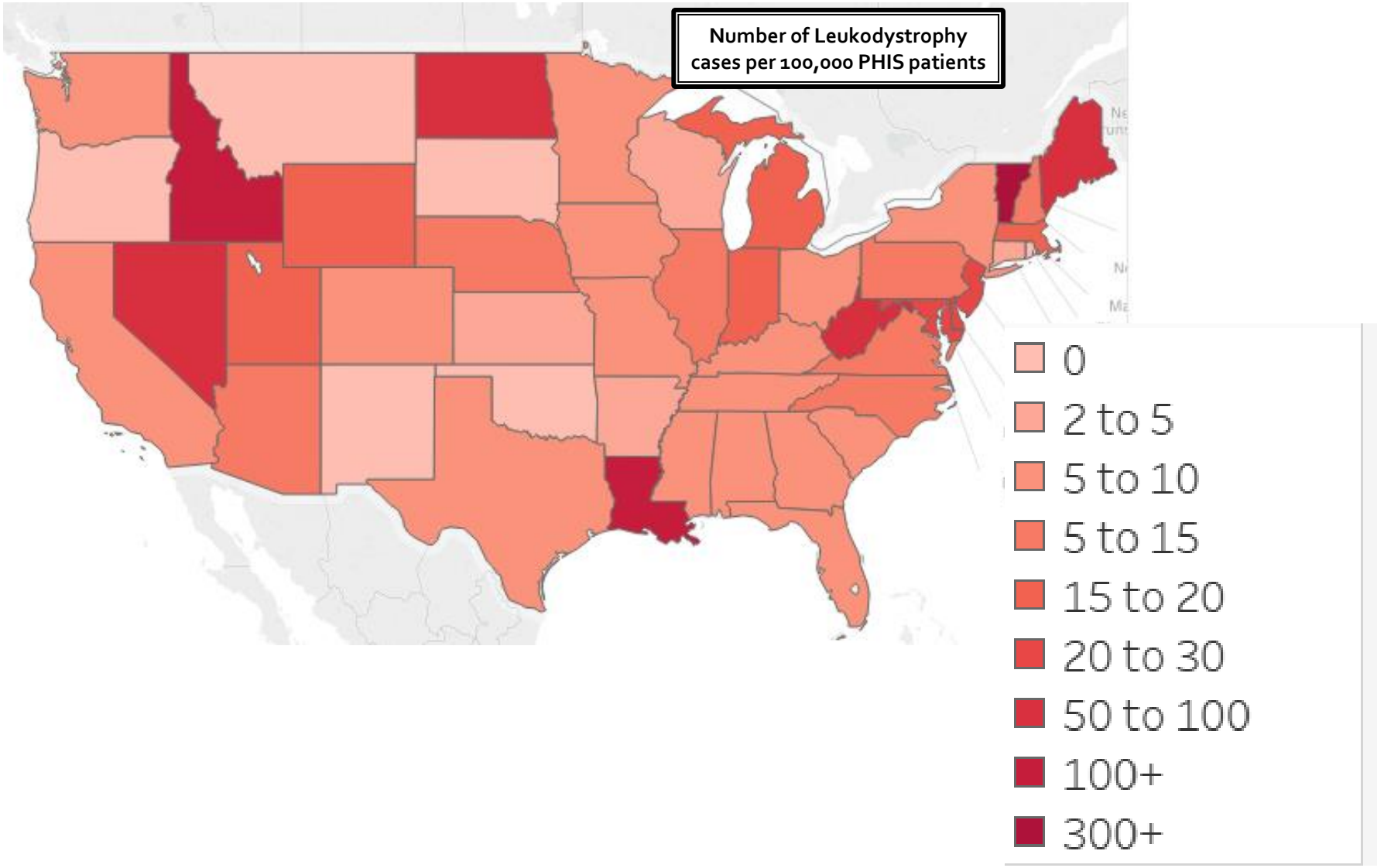
Leukocyte Lysosomal Enzymes and Serum Very Long Chain Fatty Acids

or

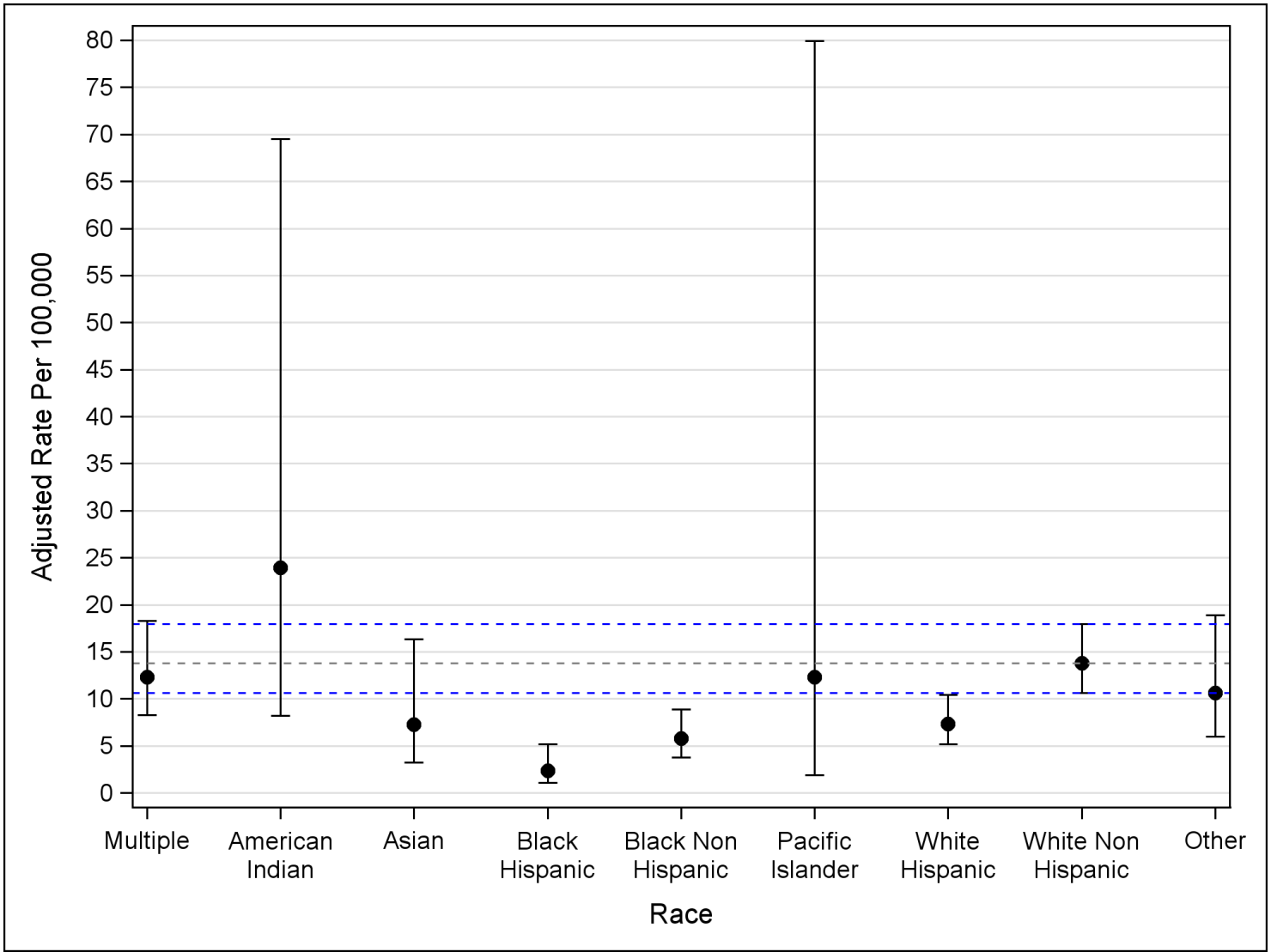
Rapid Whole Exome

2. **Whole exome/genome or leukodystrophy gene panel**

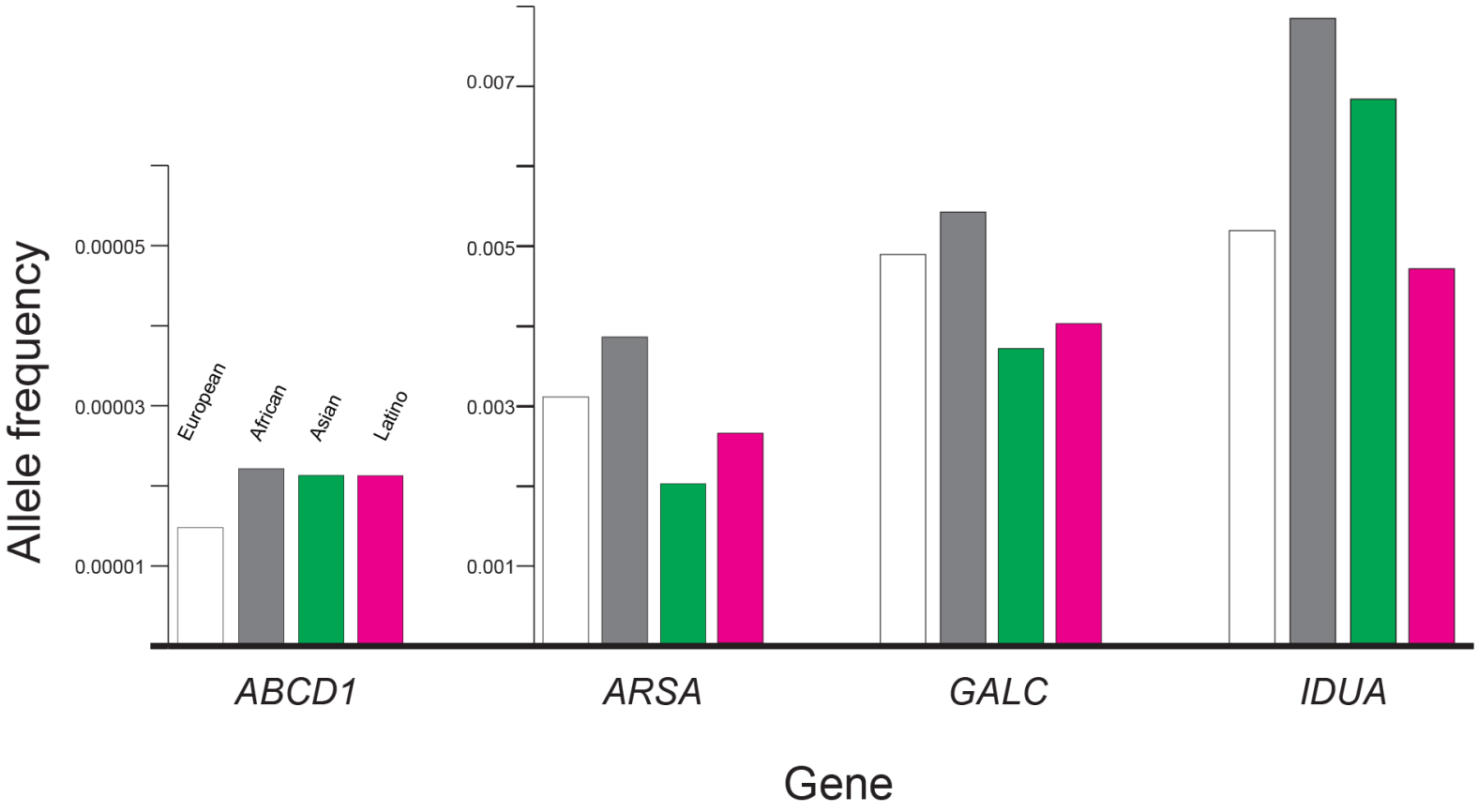
Diagnosis Disparities- a role for NGS?


Number of Leukodystrophy cases per 100,000 PHIS patients
0
2 to 5
5 to 10
5 to 15
15 to 20
20 to 30
50 to 100
100+
300+

Diagnosis rates are >50% lower in some racial groups



No evidence for genetic



Conclusions

- NGS diagnosis is less expensive
  - Than traditional diagnosis
  - Than clinical care
    - The Diagnostic Odyssey can be finite
- NGS algorithms for diagnosis should be developed
- Consider NGS to reduce diagnosis disparities

Crispy Zebrafish
(CRISPR and Zebrafish)

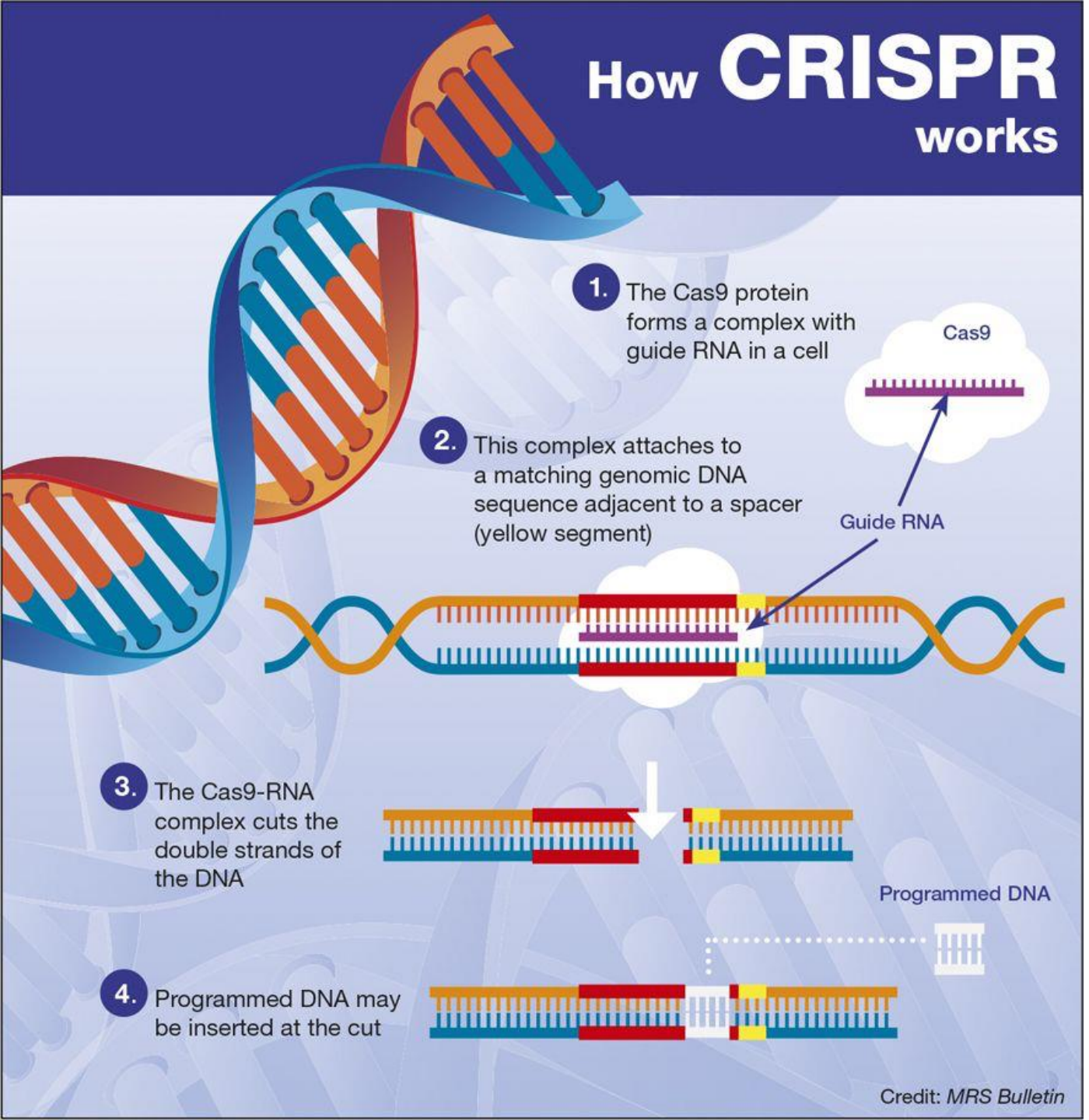
CRISPR

- CRISPR is the most recent and most successful of genome editing techniques
  - ZFN (zinc-finger nucleases)
  - TALENs (transcription activator-like effector nucleases)
  - ZFNs and TALENs require customization to efficiently target a sequence, and are more costly and difficult to develop for each target

CRISPR mechanics

- CRISPR/Cas system is a prokaryotic (bacterial) "immune" system to attack foreign DNA
  - CRISPR:
    Clustered Regularly Interspaced Short Palindromic Repeats
  - Cas: CRISPR-associated system
    - Cas9: an RNA-guided DNA endonuclease
- Synthetic gRNA (guide RNA) matches a sequence in the target, and then guides the Cas9 system over to cut at that locus

How CRISPR works

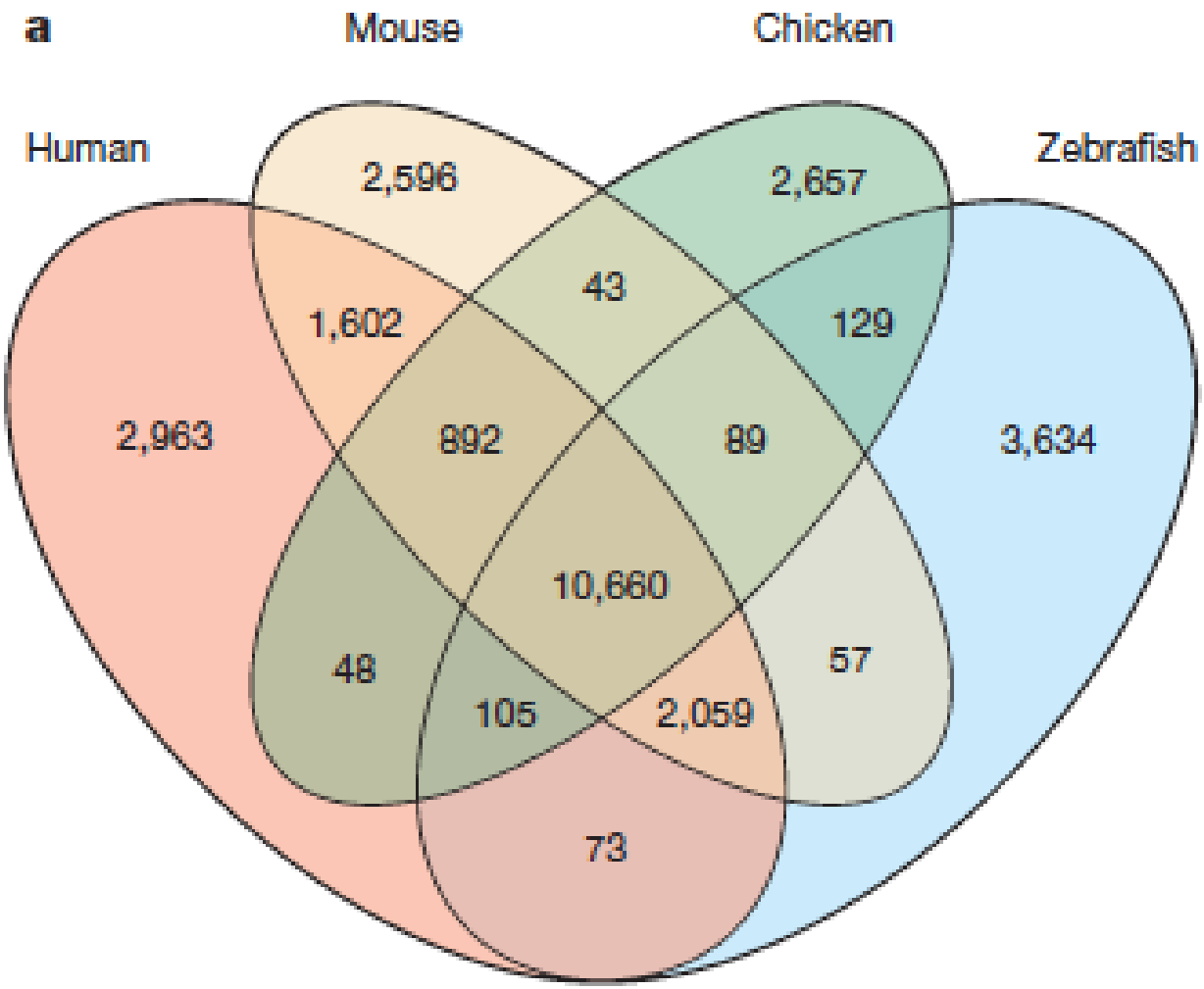


Credit: *MRS Bulletin*

Zebrafish as a Model Organism

1. Vertebrate
2. Conserved genes
3. Rapid development
4. Inexpensive

Zebrafish and Human Genes are Conserved



Using Model Systems

PNAS

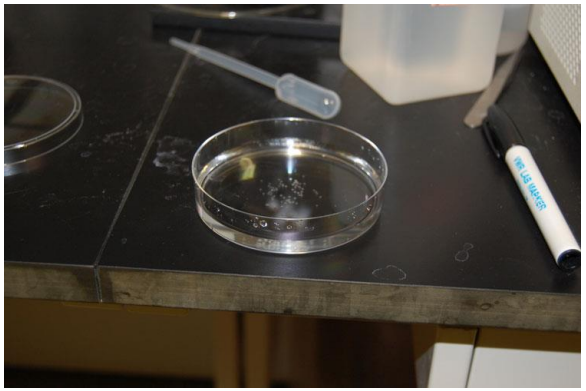
Genomic responses in mouse models poorly mimic human inflammatory diseases

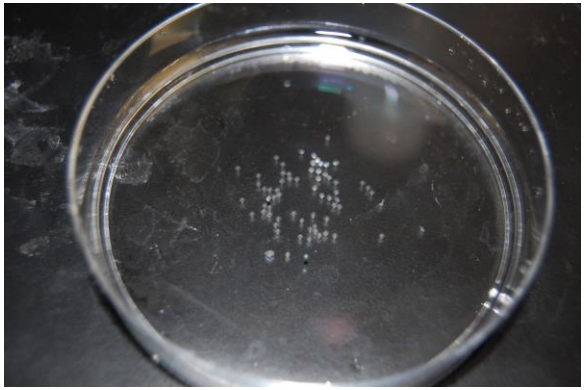
Junhee Seok[a,1], H. Shaw Warren[b,1], Alex G. Cuenca[c,1], Michael N. Mindrinos[a], Henry V. Baker[c], Weihong Xu[a], Daniel R. Richards[d], Grace P. McDonald-Smith[e], Hong Gao[a], Laura Hennessy[f], Celeste C. Finnerty[g], Cecilia M. López[c], Shari Honari[f], Ernest E. Moore[h], Joseph P. Minei[i], Joseph Cuschieri[j], Paul E. Bankey[k], Jeffrey L. Johnson[h], Jason Sperry[l], Avery B. Nathens[m], Timothy R. Billiar[l], Michael A. West[n], Marc G. Jeschke[o], Matthew B. Klein[j], Richard L. Gamelli[p], Nicole S. Gibran[j], Bernard H. Brownstein[q], Carol Miller-Graziano[k], Steve E. Calvano[r], Philip H. Mason[e], J. Perren Cobb[s], Laurence G. Rahme[t], Stephen F. Lowry[r,2], Ronald V. Maier[j], Lyle L. Moldawer[c], David N. Herndon[g], Ronald W. Davis[a,3], Wenzhong Xiao[a,t,3], Ronald G. Tompkins[t,3], and the Inflammation and Host Response to Injury, Large Scale Collaborative Research Program[4]

Among genes changed significantly in humans, the murine orthologs are close to random in matching their human counterparts (e.g., R^2 between 0.0 and 0.1)

Economy of scale

- Analyze *1000s* of animals per day
- *1000s* of tanks in a facility
- Generation time: 8 weeks

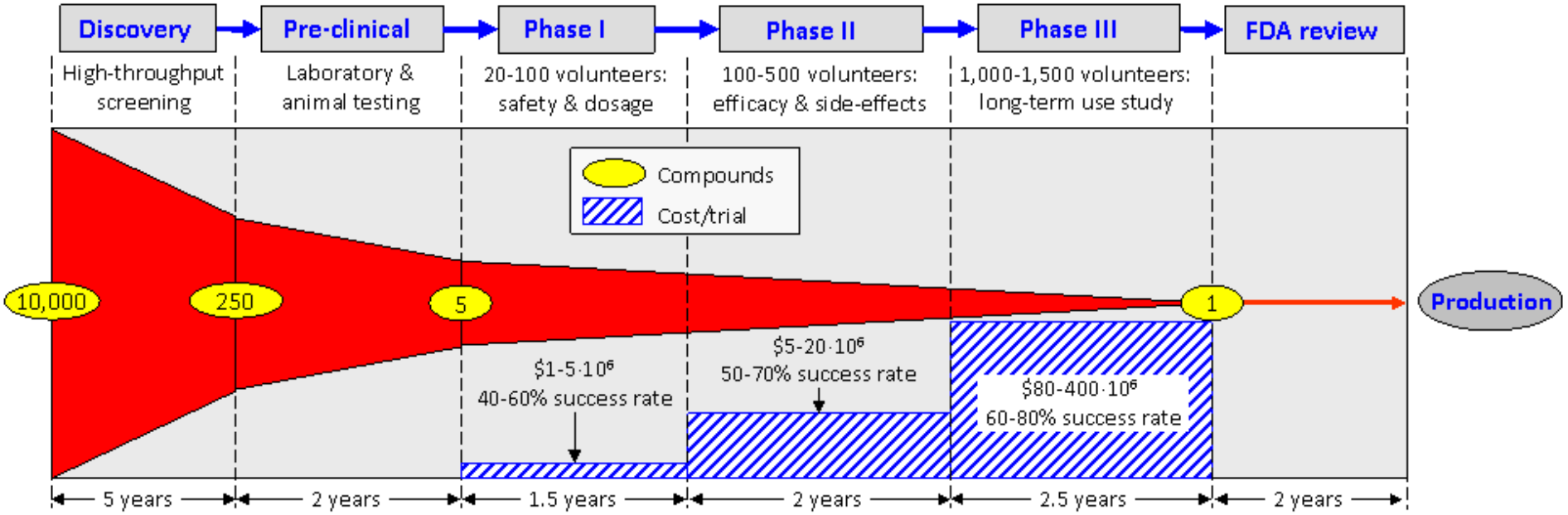




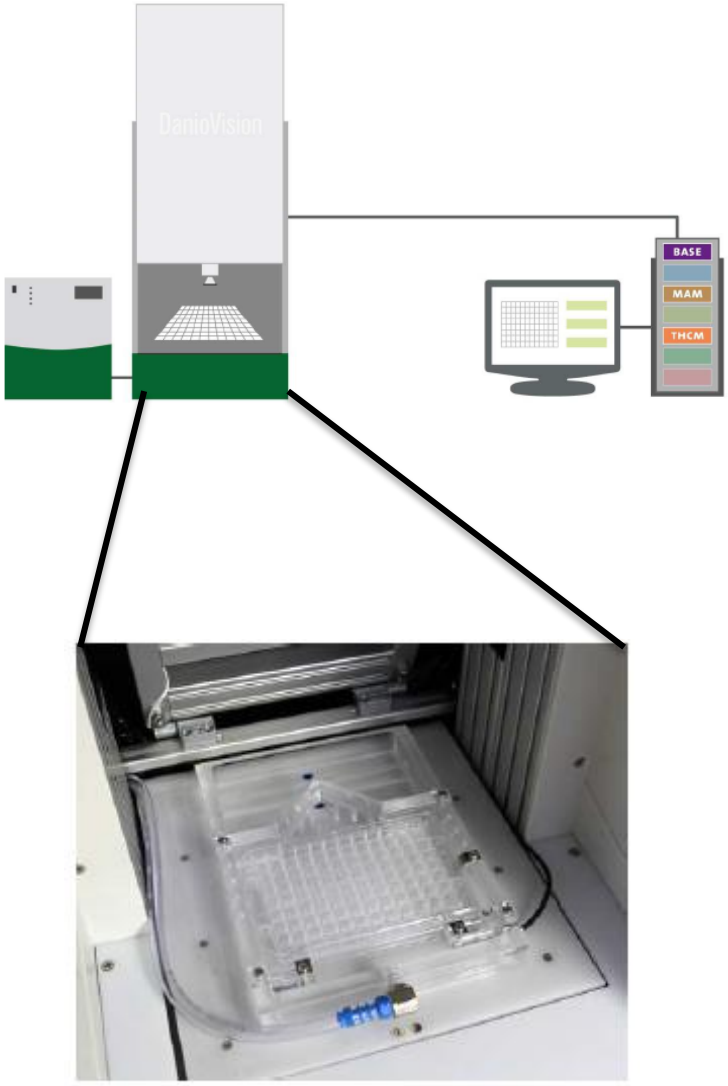
Power of Drug Discovery in Zebrafish

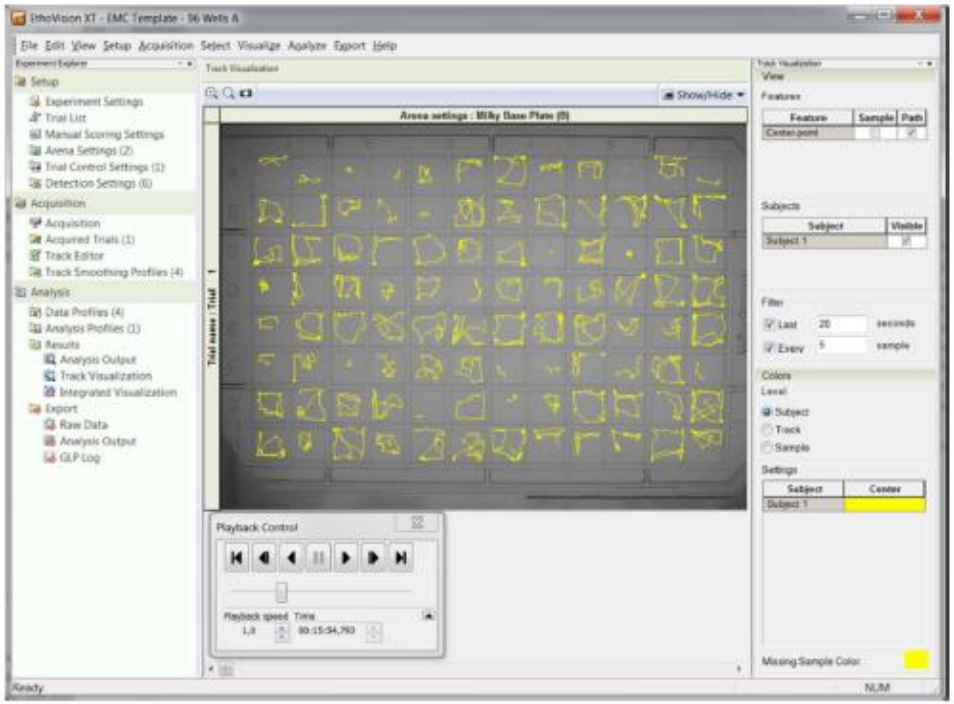
- whole animal biological complexity
- rapid development
- high-throughput screening
  - 62% of new drugs discovered using phenotypic screening

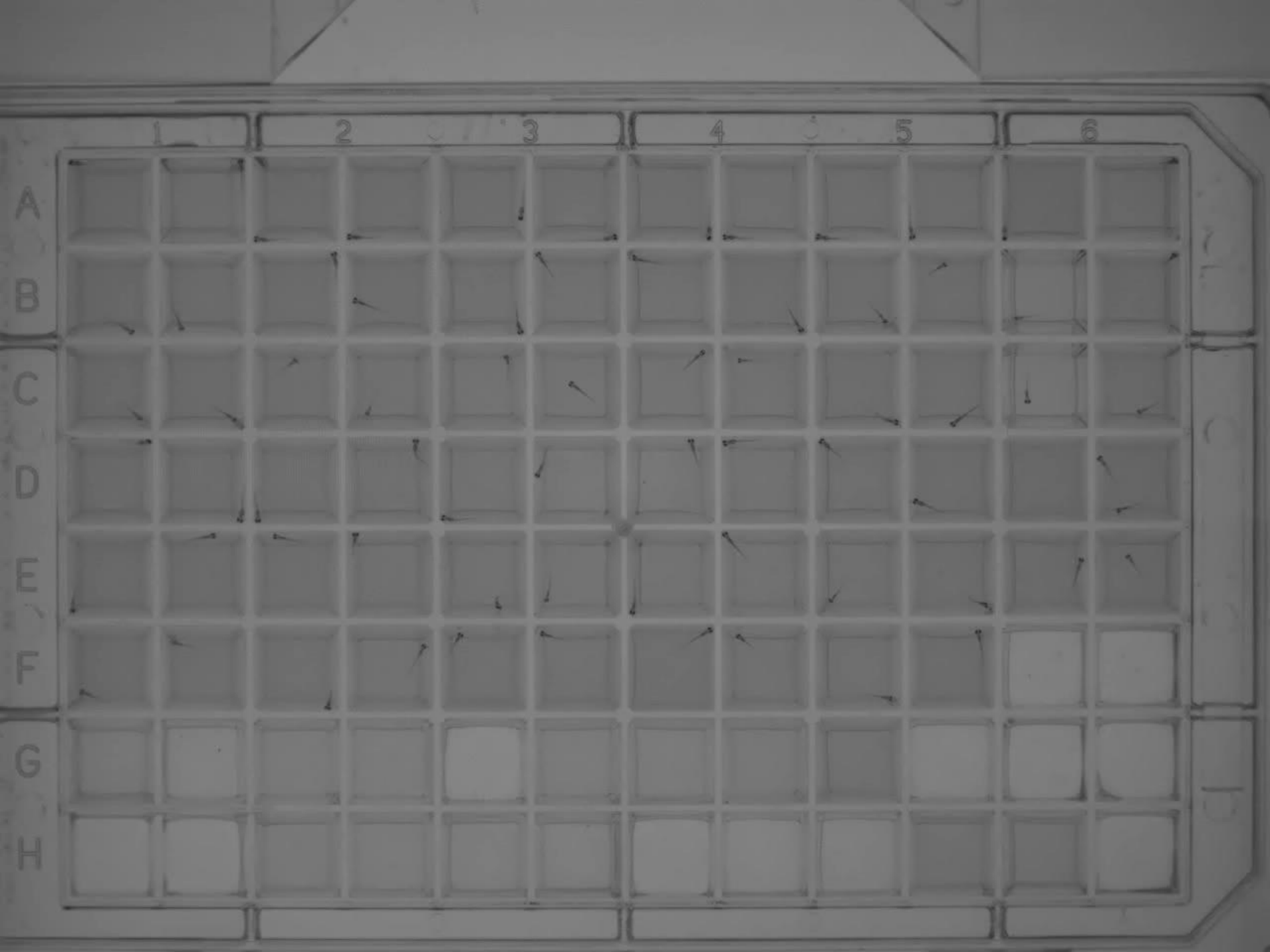
Drug Pipeline



Automated screening






1
2
3
4
5
6
A
B
C
D
E
F
G
H

CRISPR in Zebrafish

- Bi-allelic knockdown using CRISPR >80%
  - Both copies of a gene are mutated
  - From the 1-2 cell stage of life
- CRISPR construct is easy to make and can be ready in <1 week and <$400
- Multiple genes can be targeted simultaneously
- >1000 animals can be generated in a week and tested by an undergraduate
- Results can be known in 1-2 weeks for developmental disorders
  - Because embryogenesis occurs in first 3 – 7 days

Zebrafish CRISPR limits

- *Limits*
  - Some genes in the zebrafish genome are duplicated
  - A stable mutant for long-term studies takes 1 year to generate
  - Some disorders are not amenable for zebrafish (for example, thumb development, or disorders of the placenta, etc.)
  - Some "rescue" may occur by orthologs

Conclusions

Zebrafish have unique benefits as a vertebrate model organism

- rapid generation time, high numbers, and inexpensiveness

CRISPR is fast and efficient in zebrafish

Zebrafish have emerged as a powerful tool for testing NGS results

Perils and Successes with CRISPR Modeling

Three tales (tails?)

Guidelines for Demonstrating Variant Pathogenicity

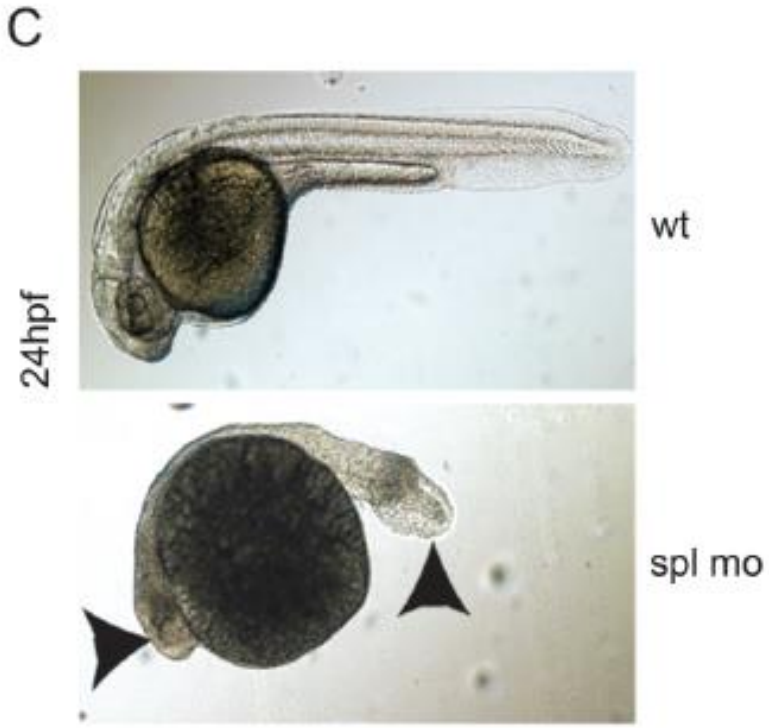
1. specific gene variant enriched/specifically associated with a disease
2. a mutant phenotype in a model system matches a phenotype from human
3. Rescue of the mutant phenotype with wild-type allele
4. Inability of mutant allele to rescue phenotype

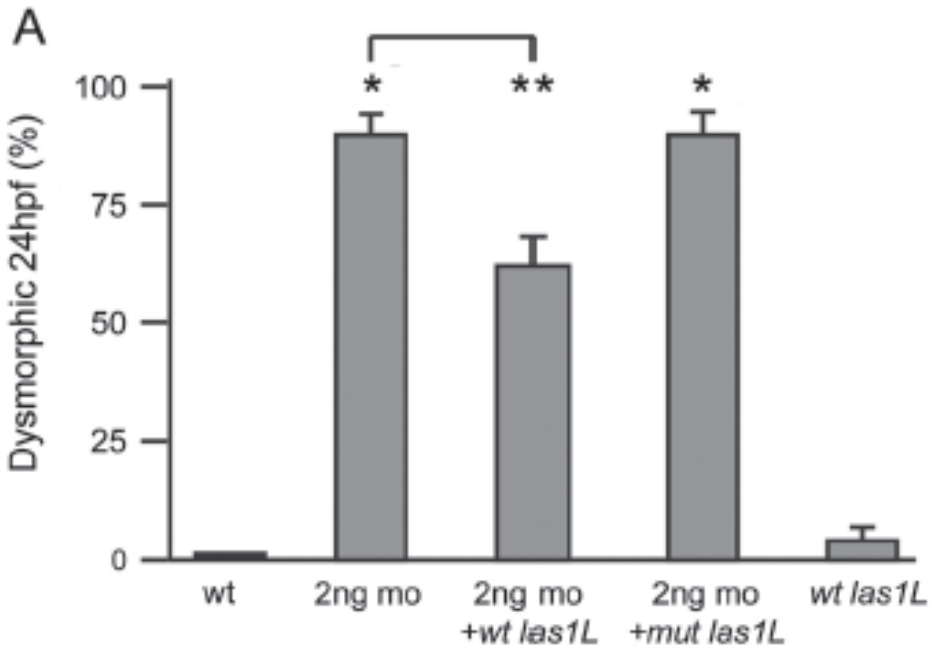
adapted from Chakravarti et al., 2013, Cell

Two congenital motor neuron diseases: …a New Gene… and the Wrong Gene

- Case 1:
  - Newborn infant requiring artificial ventilation
  - Genetic testing showed that it was not SMA
  - Guidance needed for parents and physicians

LAS1L gene identified and had phenotype in zebrafish



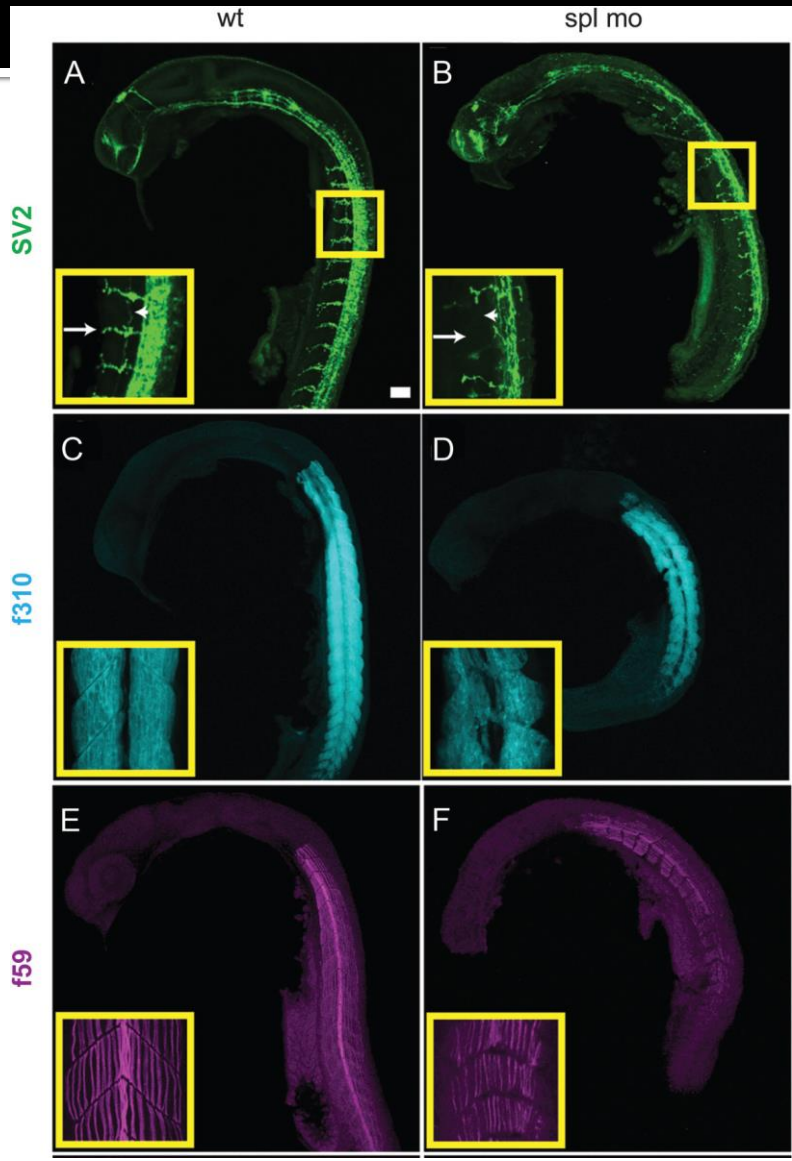


LAS1L Pathogenicity

- Sequencing showed p.S477N mutation in a ribosomal biogenesis protein: LAS1-like
- Confirmed in zebrafish
- New biochemical pathway in neurological disease

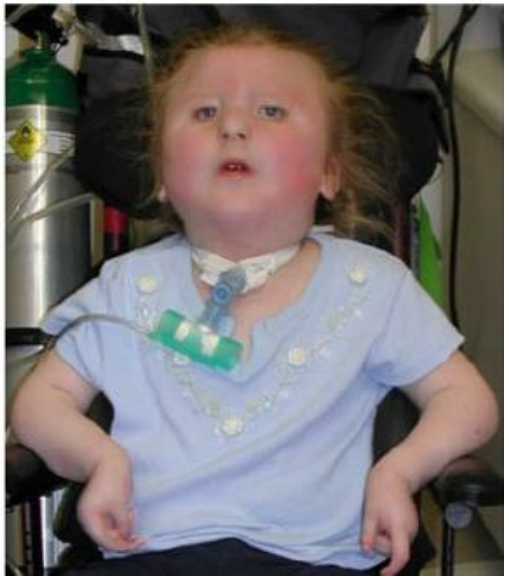
Congenital lethal motor neuron disease with a novel defect in ribosome biogenesis

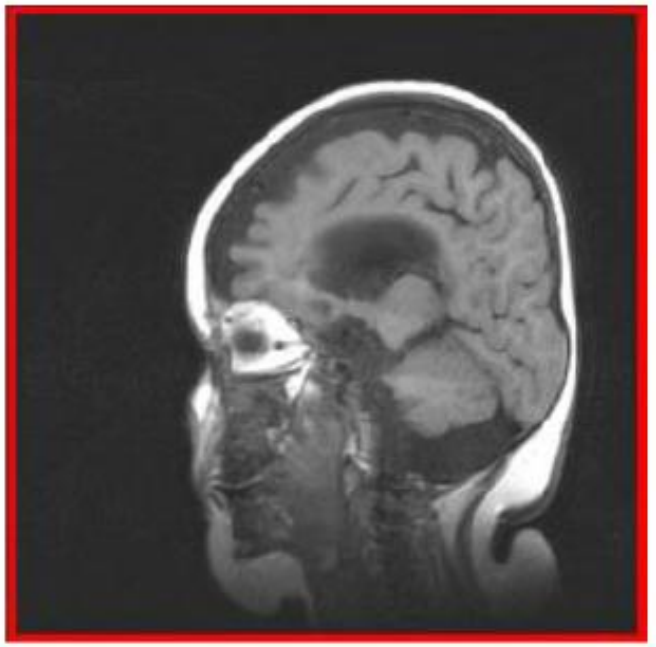
Butterfield et al., Neurology, 2014



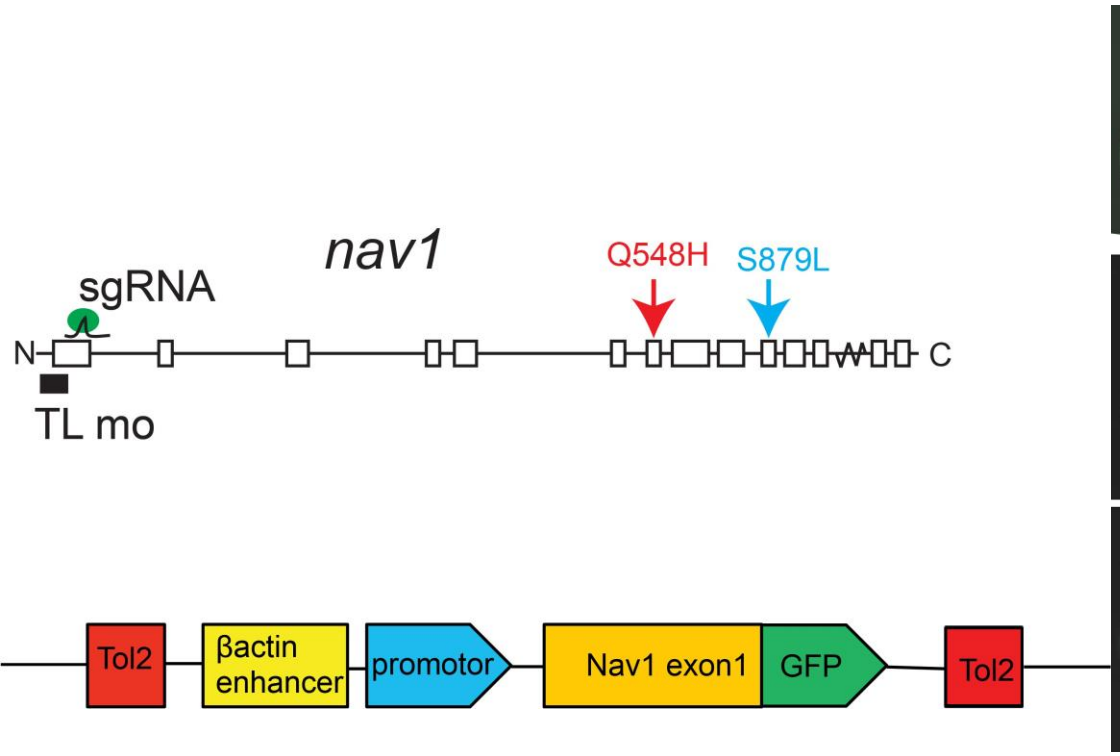
Case 2

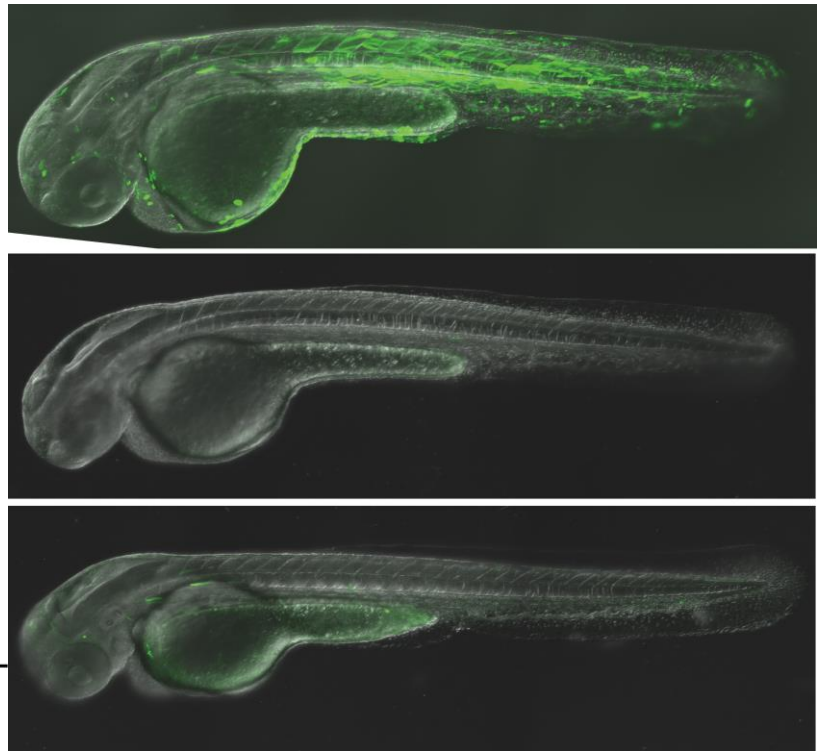
- Stevenson and Carey, AJMG, 2007
- Siblings with muscular contractures, seizures, and brain structural abnormalities
- NGS suggested NAV1 gene





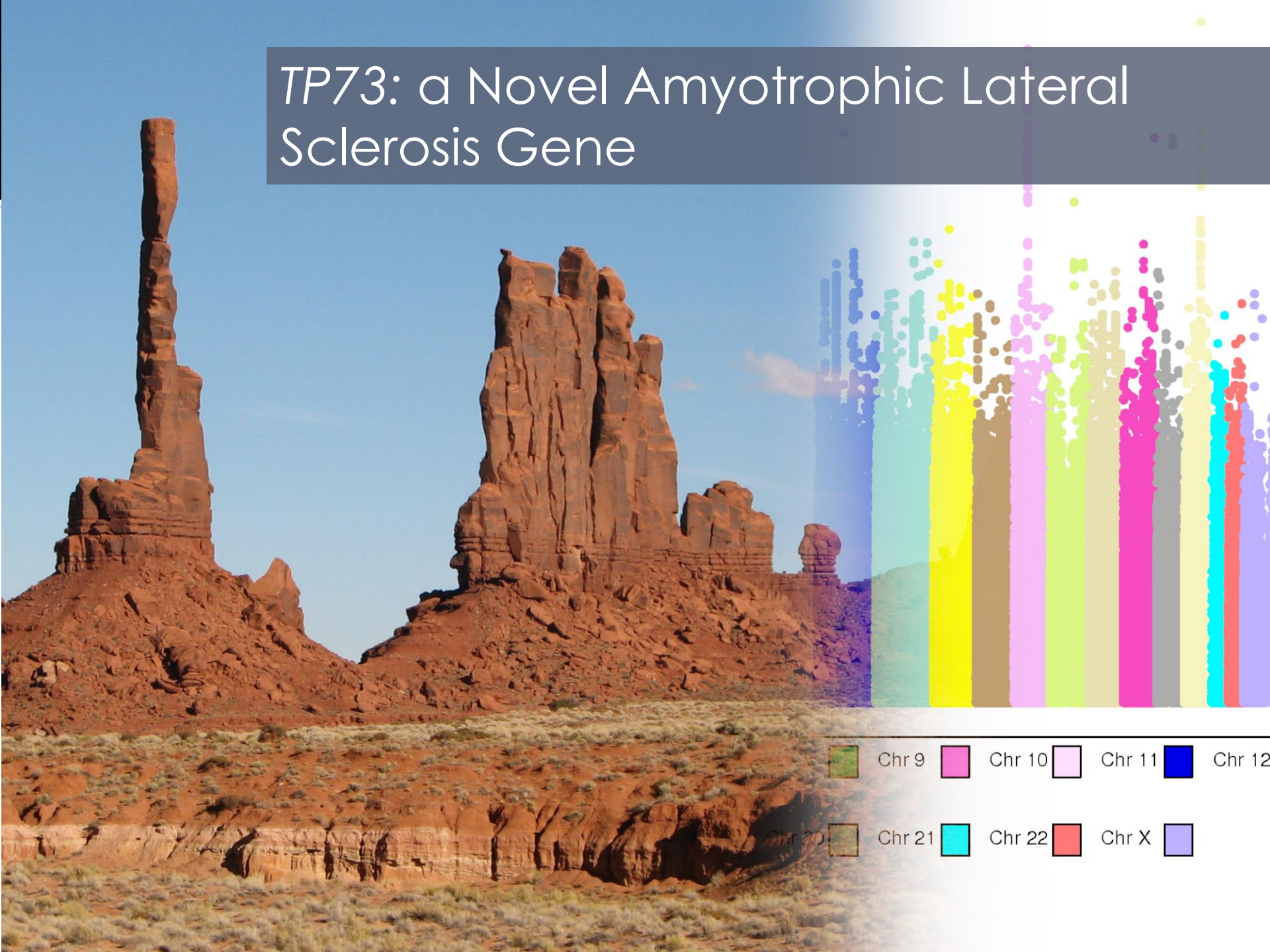
NAV1 gene incorrect





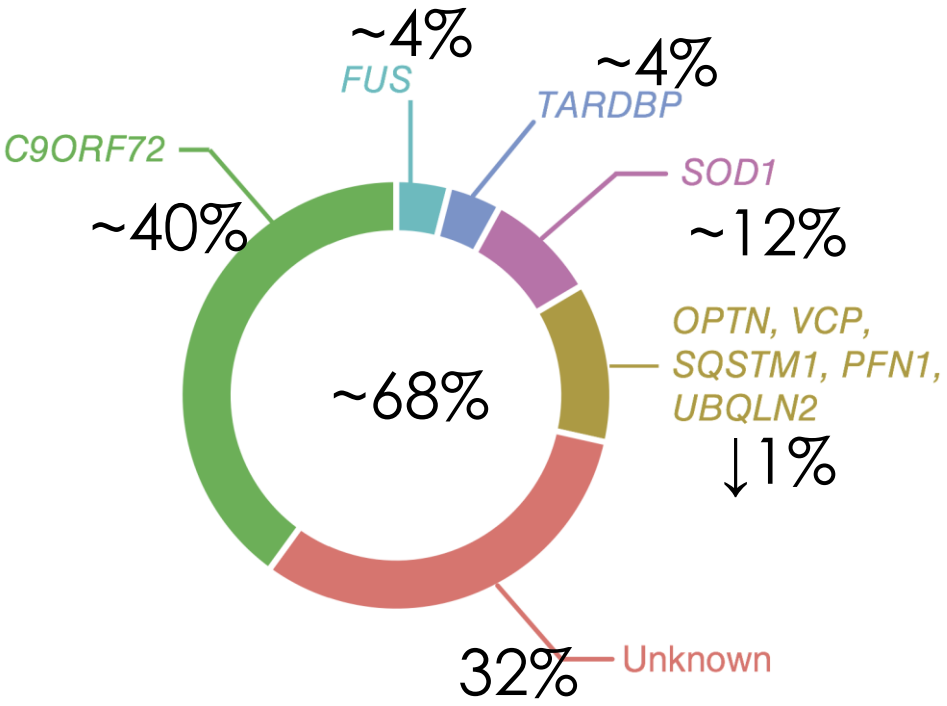
- zebrafish morphants and CRISPR are normal
  - *sequence re-analysis did not confirm NAV1 (and did not identify other better candidates)!*

TP73: a Novel Amyotrophic Lateral Sclerosis Gene



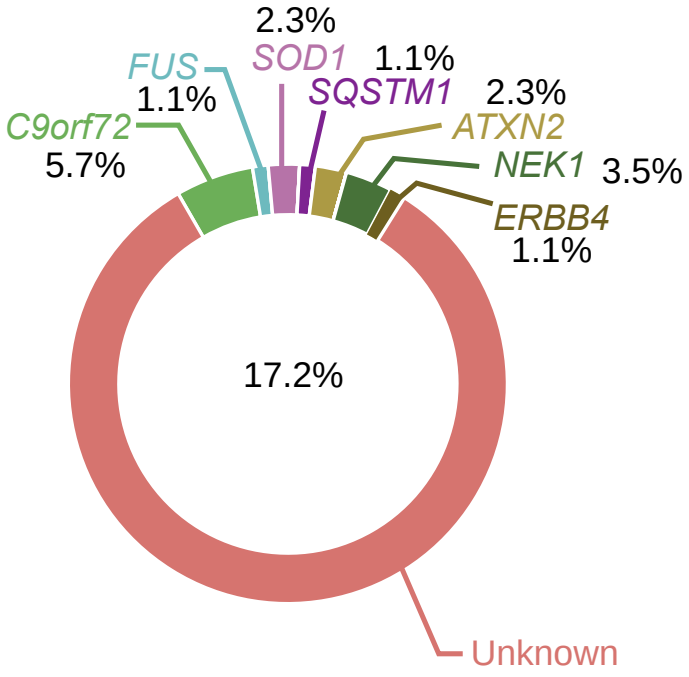
Most ALS cases have an unknown genetic cause for disease

Familial ALS (10%)



Renton et al. (2014),
Nat. Neurosci

Sporadic ALS (90%)



Gibson, Downie et al. (2017),
Neurology

Determine whether novel loci ALS loci can be identified using next-generation sequencing

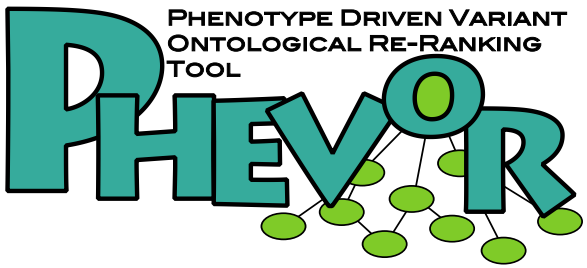
87 SALS patients
(exome sequenced)

324 controls
(Simons Simplex Collection)

Burden testing

Prioritized gene list by burden

Candidate gene list applicable to a phenotype/disease



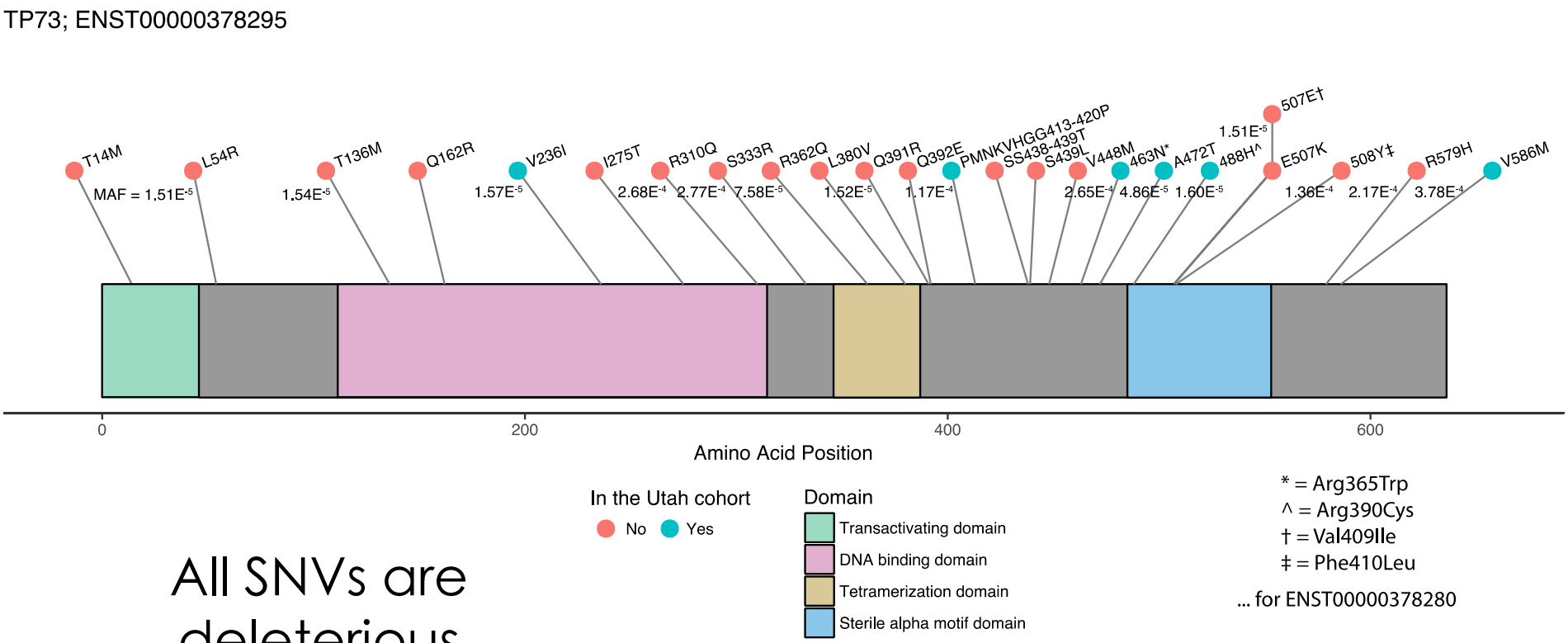
Re-ranked gene list with genes relevant to a phenotype ranked higher

TP73 has multiple qualities that make it an attractive ALS gene candidate

- Two known ALS genes in top 5 ranked genes from VAAST/PHEVOR
  - *MAPT* (rank: 3)
  - *SOD1* (rank: 5)
- *TP73* (rank 2)
  - One of two genes that possessed a VAAST burden level approaching genome-wide significance
  - Four different rare missense SNVs in five patients
    - 1 in-frame indel upon screening for indels
  - Part of the p53 family of tumor suppressor proteins
  - Neuronal survival factor

Rare, deleterious variants in *TP73* are found at appreciable frequency in ALS patients

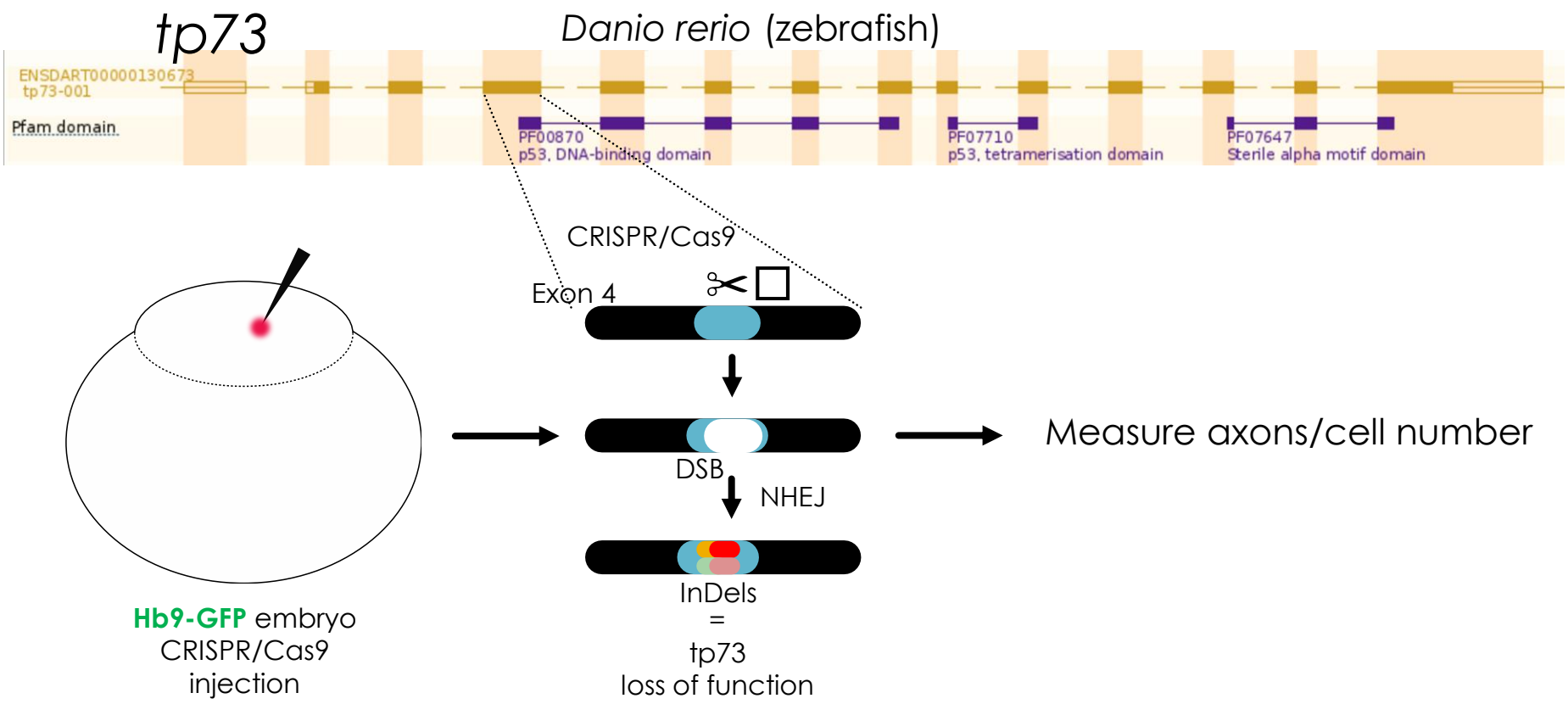
24 rare (MAF<0.0005) *TP73* coding variants were found in ~2,900 ALS patients



All SNVs are deleterious according to MetaSVM

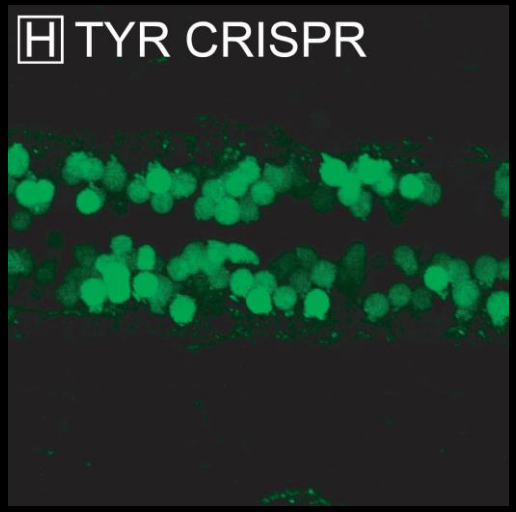
~2,800 patients from Cirulli et al. (2015) Science

A CRISPR/Cas9 zebrafish system was developed to test how loss of p73 affects spinal motor neurons

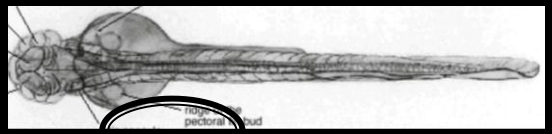


Orange = target sequence

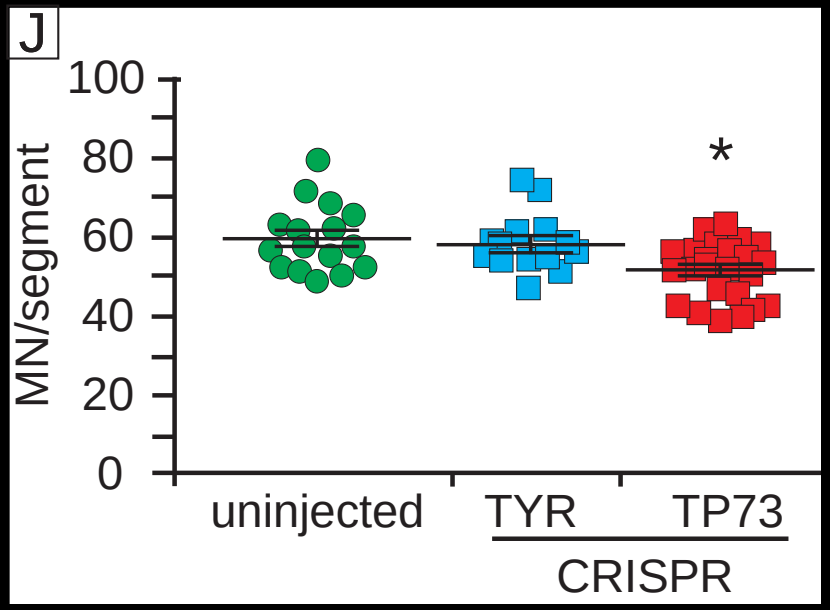
The number of spinal motor neurons is significantly reduced in *tp73* zebrafish mutants



Tg[*Hb9:Gal4-UAS:GFP*]

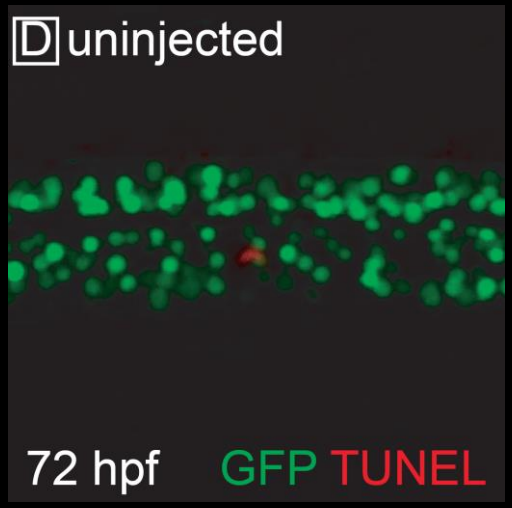


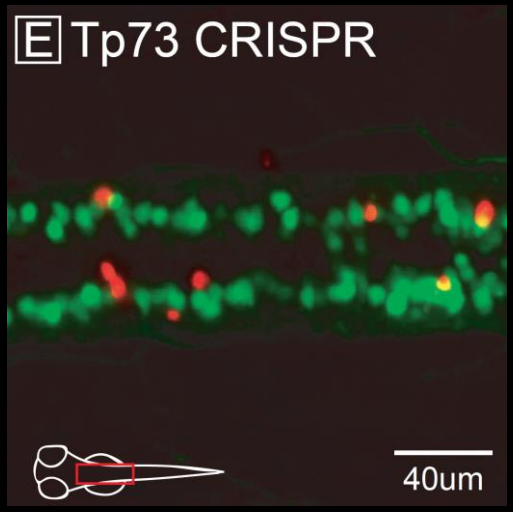
Confocal: 10x; 5µm/step, 21 steps



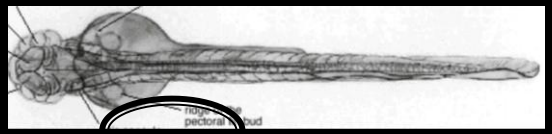
Hb9 = motor neuron promoter
hpf = hours post fertilization
MN = motor neuron
* = $p < 0.01$

p73 CRISPR zebrafish have increased apoptosis of spinal motor neurons



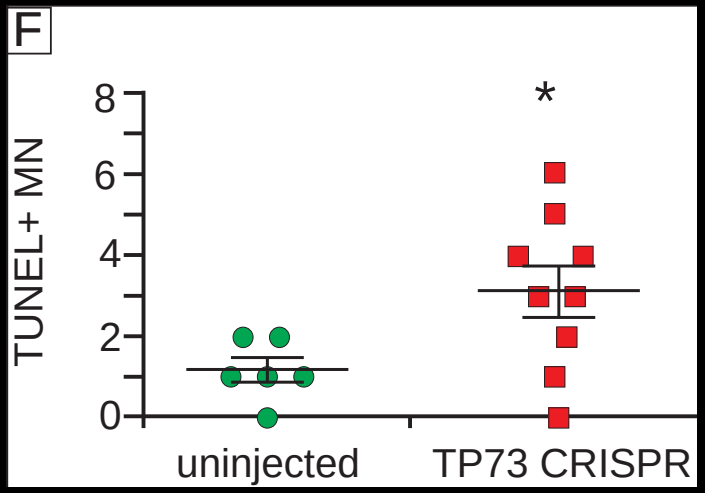


Tg[*Hb9:Gal4-UAS:GFP*]



Confocal: 10x; 5µm/step, 21 steps

Hb9 = motor neuron promoter
hpf = hours post fertilization
MN = motor neuron
* = $p < 0.05$



Conclusions

- May have identified a new ALS risk gene.
  - Rare and deleterious variants *TP73* are found in ALS patients
  - These variants impair *TP73* function
    - Loss of C2C12 myoblast ability to escape differentiation
  - Development and survival of motor neurons are negatively affected in *Tp73* mutant zebrafish
- Expands the list of cellular processes involved in ALS pathogenesis.

Acknowledgements

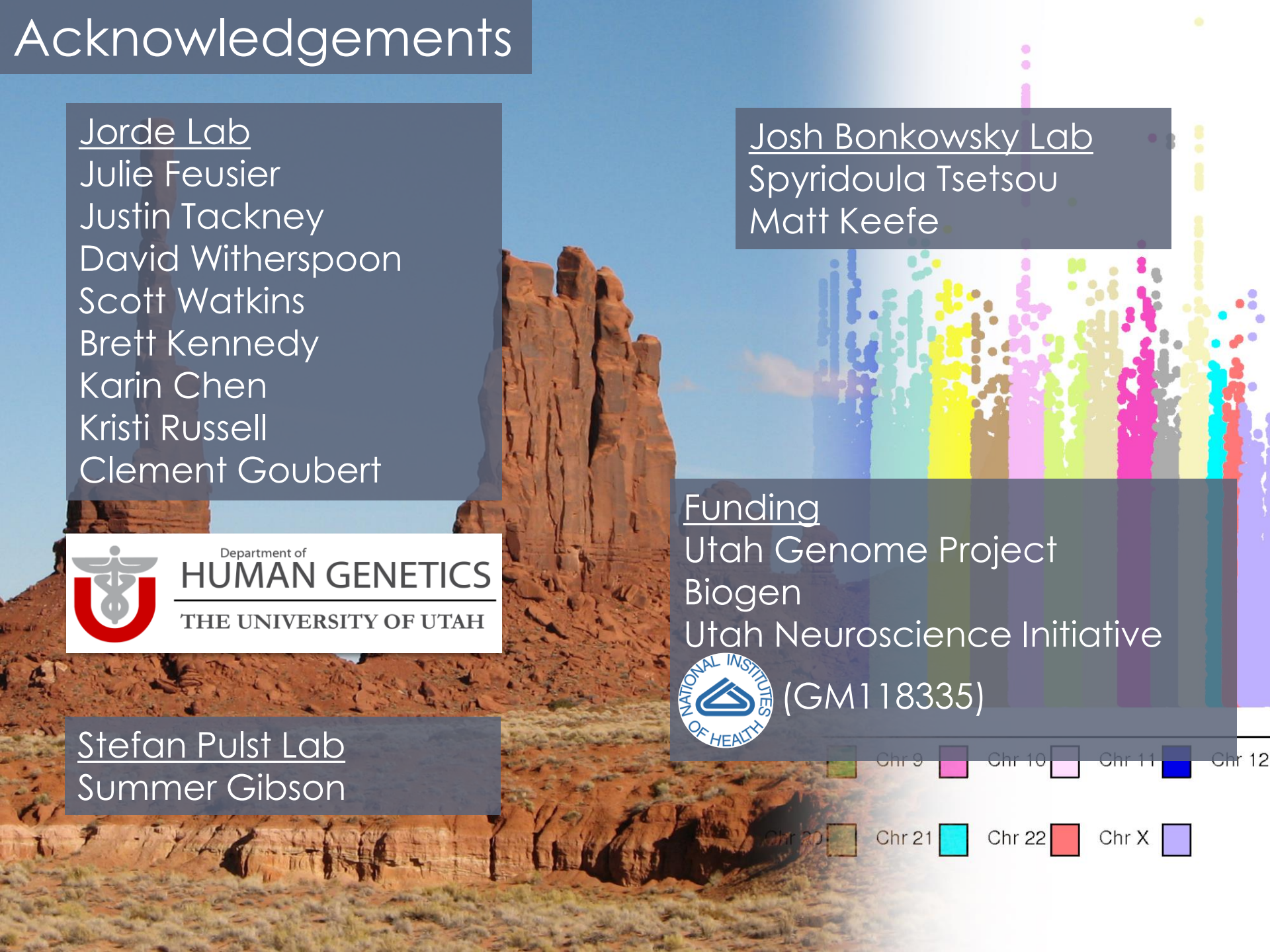
Jorde Lab
Julie Feusier
Justin Tackney
David Witherspoon
Scott Watkins
Brett Kennedy
Karin Chen
Kristi Russell
Clement Goubert

Department of HUMAN GENETICS
THE UNIVERSITY OF UTAH

Stefan Pulst Lab
Summer Gibson

Josh Bonkowsky Lab
Spyridoula Tsetsou
Matt Keefe

Funding
Utah Genome Project
Biogen
Utah Neuroscience Initiative
NATIONAL INSTITUTES OF HEALTH (GM118335)



Conclusions

- NGS is changing the landscape not only of diagnosis, but redefining what diseases exist
- NGS results can be challenging to interpret, as often the results are the first of their kind
- CRISPR genome editing is a powerful, efficient, and inexpensive method for testing gene function
- The zebrafish is a uniquely powerful vertebrate model system for testing certain diseases and NGS results

Funding

THE Vanishing White Matter Foundation

People

University Health Care
Pediatrics