Rare Disease-Target Gene Treatment Knowledge Mining Partnership between Rancho BioSciences (Rancho) and Rady Children's Institute for Genomic Medicine (RCIGM)

Christine Marciniak[1], Anna Chernysheva[1], Akshata Varun[1], Guang-dan Zhu[1], Laurie D. Smith[2], Sara Caylor[2,3], David Dimmock[2], Annette Feigenbaum[2,3,4], Alena Fedarovich[1], Christian Hansen[2,3], Daeheon Oh[2,3], Elena Pahomova[1], Liana Protopsaltis[2,3], Rebecca Reimers[2,3], Tatiana Sitnik[1], Lauren Stoddard[2,3], Kristen Wigby[2,3], Meredith Wright[2,3], and Stephen F. Kingsmore[2,3,5]

[1]Rancho BioSciences, PO Box 7208, Rancho Santa Fe, CA 92067; [2]Rady Children's Institute for Genomic Medicine, San Diego CA 92123 USA; [3]Rady Children's Hospital, San Diego CA 92123, USA, [4]Department of Pediatrics, University of California San Diego, San Diego CA 92093 USA; [5]Keck Graduate Institute, Claremont, CA 91711 USA

Combined Totals to Date

- 1040 Disease-Gene Pairs Curated
- 18003 Treatments Identified
- 2043 Treatments Retained After Review (11.3%)

Abstract

Annually, up to 10% of the nearly 3.6 million infants born in the US are admitted to a hospital, many of which end up in an intensive care unit (ICU). About 33% of those infants hospitalized in an ICU are thought to have an underlying rare genetic disease and may be a candidate for individualized therapy. Delays in time to intervention occur when a provider spends hours pulling and reviewing relevant literature or when access to specialty care is unavailable.

RCIGM has developed a Genome-to-Treatment platform in conjunction with Rancho to provide accurate information when a rare disorder is identified to enable timely informed decisions about treatment.

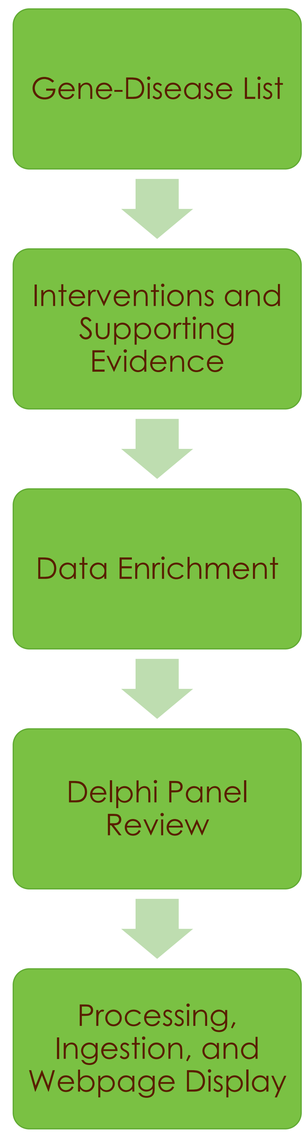
Knowledge mining and data curation from public data sources was performed by Rancho to find existing treatments associated with any variants of a target gene and disease of interest. Originally, public data sources were manually searched and curated for all interventions mentioned. Automated processes now increase data acquisition and reduce time to results. 700 gene-disease combinations and thousands of public data sources have been reviewed and curated. A combined governance team assess the complexity of work and progress against disease-target gene coverage. Focus on pediatric cases and most recent evidence from human studies resulted in over 10,000 potential interventions submitted for review.

After data mining for gene-disease-treatment was completed, the information is reviewed by a dedicated clinical and biochemical genetics team to identify 1) if the condition is appropriate for inclusion in a rapid-whole-genome newborn screening panel and 2) what interventions are appropriate for inclusion in a web-based tool, either for acute care or for newborn screening follow-up (Gene-to-Treatment; BeginNGS). Data is collected using REDCap. Additional review is performed by a physician and laboratory director comprised workgroup to evaluate final inclusion to BeginNGS.

585 disorders were reviewed for ICU presentations and 455 disorders reviewed for newborn screening. Of the interventions that were curated, 1,208/10,380 (11.6%) were retained in the ICU set and 835/7,623 (11.0%) were retained in the NBS dataset. Through the course of this partnership, time per gene-disease pair has decreased from >15 hours in 2020 to <5 hours.

Leveraging advances in technology for information curation of gene-disease-treatment is increasingly important as more infants are screened and the number of available treatments expands.

Workflow



Interventions:

- Focus on pediatric treatments
- Focus on those that stabilize/improve, though contraindications included
- Categories: medicine, surgery, diet, device, other
- 5-10 most recent years

Sources:

- PubMed, Conference abstracts
- Data aggregation sites
- Clinical trials

Enriched with:

- OMIM/Orphanet IDs
- URLs
- GARD, GHR, DrugBank
- Incidence Data
- Age of Onset (neonate, infant, child)
- Clinical Summaries

Delphi Panel Review Criteria:

- Efficacy
- Appropriateness to pediatric setting
- Precautions evaluated
- Subpopulations of affected cases
- Categories: Curative, Effective/Ameliorative, Still in Trials/Unproven, Contraindicated
- Rated on priority of use (first, second, third line treatment, etc.)
- Estimate need to treat the disorder (urgent, semi-urgent, not urgent)
- Quality of supporting evidence (Authoritative published clinical practice guideline, Cohort Study or Studies, Case Report(s), Expert Opinion)

Methodology Overview

- Knowledge Mining/Data Integrity and QC
  - Evidence researched for diseases with existing treatments that may be associated with any variants of the genes
  - Template populated and QCed by a second independent scientist
- Governance
  - Complexity of the work and curation progress against the target gene-disease coverage assessed
- RedCap processing
  - Reviewers record edits and/or evaluation, full team meet and compare for final judgement and decision; revisited annually for updates/revisions

Selected Sites Using GTRx

Balboa Naval Medical Center
Brawley Hospital
Camp Pendleton Naval Hospital
Grossmont Hospital
Inland Urgent Care
Kaiser Permanente
Loma Linda Medical Center
Naval Hospital Camp Pendleton
Palomar
Pioneers Memorial
Rancho Spring Medical Center
RCHSD
Rutgers U Medical Center
Scripps Medical Network
Sharp Medical Network
U of Tennessee
UCSD Medical Network
U of Colorado Anschutz Medical Campus

Disease-Gene-Treatment Statistics

Disease Name	Gene	# treatments curated	# treatments approved and retained	% treatments approved and retained
Abetalipoproteinemia	*MTTP*	16	3	18.8%
Adams-Oliver syndrome 5	*NOTCH1*	2	1	50.0%
Adenine phosphoribosyltransferase deficiency	*APRT*	8	4	50.0%
ADRENAL HYPERPLASIA, CONGENITAL, DUE TO 21-HYDROXYLASE DEFICIENCY*	***CYP21A2***	**8**	**2**	**25.0%**
Adrenal hypoplasia, congenital	*NR0B1*	35	4	11.4%
AGAMMAGLOBULINEMIA 8A, AUTOSOMAL DOMINANT	*TCF3*	4	1	25.0%
AGAMMAGLOBULINEMIA 9, AUTOSOMAL RECESSIVE	*SLC39A7*	7	2	28.6%
ANEMIA, NONSPHEROCYTIC HEMOLYTIC, DUE TO G6PD DEFICIENCY*	***G6PD***	**21**	**5**	**23.8%**
ANEMIA, X-LINKED, WITH OR WITHOUT NEUTROPENIA AND/OR PLATELET ABNORMALITIES	*GATA1*	4	3	75.0%
ASPARTYLGLUCOSAMINURIA	*AGA*	5	3	60.0%
Ataxia with isolated vitamin E deficiency	*TTPA*	15	3	20.0%
ATAXIA-PANCYTOPENIA SYNDROME	*SAMD9L*	10	1	10.0%
Bamforth-Lazarus syndrome	*FOXE1*	27	4	14.8%
Bare lymphocyte syndrome, type II, complementation group D	*RFXAP*	7	2	28.6%
Bartter syndrome, type 1	*SLC12A1*	17	1	5.9%
Bernard Soulier syndrome	*GP1BA*	3	2	66.7%
Bile acid conjugation defect 1	*BAAT*	13	1	7.7%
Blackfan-Diamond anemia*	***GATA1***	**4**	**3**	**75.0%**
CINCA SYNDROME	*NLRP3*	13	1	7.7%
Congenital erythropoietic porphyria	*GATA1*	69	1	1.4%
CYSTIC FIBROSIS*	***CFTR***	**82**	**1**	**1.2%**
ECTODERMAL DYSPLASIA AND IMMUNODEFICIENCY 1	*IKBKG*	3	2	66.7%
HYPOKALEMIC PERIODIC PARALYSIS, TYPE 2	*SCN4A*	15	2	13.3%
Hypothyroidism, congenital, nongoitrous, 8	*TBL1X*	21	3	14.3%
Infantile Sudden Cardiac Failure	*PPA2*	17	6	35.3%
Krabbe GALACTOCEREBROSIDASE (galactosylceramidase)	*GALC*	25	1	4.0%
LI-FRAUMENI SYNDROME	*TP53*	1	1	100.0%
Mannosidosis, alpha-, types I and II	*MAN2B1*	6	1	16.7%
Mitochondrial DNA depletion syndrome 2 (myopathic type)	*TK2*	72	3	4.2%
MUCOPOLYSACCHARIDOSIS, TYPE VI*	***ARSB***	**13**	**1**	**7.7%**
MUCOPOLYSACCHARIDOSIS, TYPE VII	*GUSB*	21	5	23.8%
MYASTHENIC SYNDROME, CONGENITAL, 16	*SCN4A*	16	2	12.5%
MYOTONIA, POTASSIUM-AGGRAVATED	*SCN4A*	7	3	42.9%
N-ACETYLGLUCOSAMINE-6-SULFATASE	*GNS*	18	9	50.0%
Niemann-Pick A	*SMPD1*	18	2	11.1%
NKX2-5 Congenital Heart Disease	*NKX2-5*	1	1	100.0%
Noonan syndrome 1	*PTPN11*	1	1	100.0%
NORMOKALEMIC PERIODIC PARALYSIS	*SCN4A*	14	3	21.4%
PARAMYOTONIA CONGENITA OF VON EULENBURG*	*SCN4A*	7	1	14.3%
PHENYLKETONURIA*	***PAH***	**16**	**3**	**18.8%**
Rett syndrome, congenital variant	*FOXG1*	2	1	50.0%
Silver-Russell syndrome 3	*IGF2*	37	4	10.8%
THROMBOCYTOPENIA, X-LINKED, WITH OR WITHOUT DYSERYTHROPOIETIC ANEMIA	*GATA1*	6	4	66.7%
Thyroid dyshormonogenesis 4	*IYD*	9	2	22.2%
Ventricular tachycardia, catecholaminergic polymorphic, 2	*CASQ2*	19	3	15.8%
Vitamin K-dependent clotting factors, combined deficiency of, 1	*GGCX*	14	4	28.6%

Selected examples of Disease and Gene Pairs (disease and gene names), number of interventions curated, number and percent of treatments approved and retained after review. ***Diseases identified in Phase 3 Clinical Trial.**

Impact

- Reduction in time from rare disease diagnosis to potential intervention is integral in the treatment of pediatric cases.
- A **Phase 3 Clinical Trial is ongoing at RCIGM** – so far 750 healthy children with no prior symptoms have been enrolled; preliminary results have found 28 (3.7%) positive for rare disease diagnosis after Whole Genome Sequencing. (See table for some examples of diseases identified in this study.)
- **ALL 28 positive cases identified have benefited** from this project and received interventions derived from Rancho's curation work and GTRx/BeginNGS data.
- We have not only **identified treatments available for these rare diseases** but also made access to this information through a curated, reviewed, and harmonized database ready at the physician's fingertips with **time to implementation possible in mere minutes after diagnosis**, rather than hours to days.