

# ALG1-CDG: Clinical and Molecular Characterization of 39 Unreported Patients

Bobby G. Ng,<sup>1</sup> Sergey A. Shiryayev,<sup>1</sup> Daisy Rymen,<sup>2,3</sup> Erik A. Eklund,<sup>4</sup> Kimiyo Raymond,<sup>5</sup> Martin Kircher,<sup>6</sup> Jose E. Abdenur,<sup>7,8</sup> Fusun Alehan,<sup>9</sup> Alina T. Midro,<sup>10</sup> Michael J. Bamshad,<sup>6,11</sup> Rita Barone,<sup>12</sup> Gerard T. Berry,<sup>13</sup> Jane E. Brumbaugh,<sup>14</sup> Kati J. Buckingham,<sup>11</sup> Katie Clarkson,<sup>15</sup> F. Sessions Cole,<sup>16</sup> Shawn O'Connor,<sup>16</sup> Gregory M. Cooper,<sup>17</sup> Rudy Van Coster,<sup>18</sup> Laurie A. Demmer,<sup>19</sup> Luisa Diogo,<sup>20</sup> Alexander J. Fay,<sup>21</sup> Can Ficicioglu,<sup>22</sup> Agata Fiumara,<sup>23</sup> William A. Gahl,<sup>24</sup> Rebecca Ganetzky,<sup>22</sup> Himanshu Goel,<sup>25</sup> Lyndsay A. Harshman,<sup>14</sup> Miao He,<sup>26</sup> Jaak Jaeken,<sup>3</sup> Philip M. James,<sup>27</sup> Daniel Katz,<sup>28</sup> Liesbeth Keldermans,<sup>2</sup> Maria Kibaek,<sup>29</sup> Andrew J. Kornberg,<sup>30</sup> Katherine Lachlan,<sup>31</sup> Christina Lam,<sup>32</sup> Joy Yaplito-Lee,<sup>33</sup> Deborah A. Nickerson,<sup>6</sup> Heidi L. Peters,<sup>33</sup> Valerie Race,<sup>2</sup> Luc Régál,<sup>34</sup> Jeffrey S. Rush,<sup>35</sup> S. Lane Rutledge,<sup>36</sup> Jay Shendure,<sup>6,37</sup> Erika Souche,<sup>2</sup> Susan E. Sparks,<sup>38</sup> Pamela Trapani,<sup>14</sup> Amarilis Sanchez-Valle,<sup>39</sup> Eric Vilain,<sup>40,41</sup> Arve Vølle,<sup>42</sup> Charles J. Waechter,<sup>35</sup> Raymond Y. Wang,<sup>7,8</sup> Lynne A. Wolfe,<sup>24</sup> Derek A. Wong,<sup>41</sup> Tim Wood,<sup>15</sup> Amy C. Yang,<sup>43</sup> University of Washington Center for Mendelian Genomics,<sup>1</sup> Gert Matthijs,<sup>2</sup> and Hudson H. Freeze<sup>1\*</sup>

<sup>1</sup> Human Genetics Program, Sanford Burnham Prebys Medical Discovery Institute, La Jolla, California; <sup>2</sup> Center for Human Genetics, University of Leuven, Leuven, Belgium; <sup>3</sup> Center for Metabolic Diseases, University Hospital of Leuven, Leuven, Belgium; <sup>4</sup> Section of Experimental Pediatrics, Department of Clinical Sciences, Lund University, Lund, Sweden; <sup>5</sup> Biochemical Genetics Laboratory, Mayo Clinic College of Medicine, Rochester, Minnesota; <sup>6</sup> Department of Genome Sciences, University of Washington, Seattle, Washington; <sup>7</sup> Division of Metabolic Disorders, Children's Hospital of Orange County, Orange, California; <sup>8</sup> Department of Pediatrics, University of California-Irvine School of Medicine, Orange, California; <sup>9</sup> Division of Pediatric Neurology, Baskent University School of Medicine, Ankara, Turkey; <sup>10</sup> Department of Clinical Genetics, Medical University, Bialystok, Poland; <sup>11</sup> Department of Pediatrics, University of Washington, Seattle, Washington; <sup>12</sup> Pediatric Neurology Policlinico, University of Catania, Catania, Italy; <sup>13</sup> Department of Pediatrics, Harvard Medical School, Boston, Massachusetts; <sup>14</sup> Stead Family Department of Pediatrics, University of Iowa Children's Hospital, Iowa City, Iowa; <sup>15</sup> Greenwood Genetic Center, Greenwood, South Carolina; <sup>16</sup> Division of Newborn Medicine, Department of Pediatrics, Washington University School of Medicine, St. Louis, Missouri; <sup>17</sup> HudsonAlpha Institute for Biotechnology, Huntsville, Alabama; <sup>18</sup> Department of Pediatrics, Division of Pediatric Neurology and Metabolism, University Hospital Gent, Gent, Belgium; <sup>19</sup> Clinical Genetics Program, Carolinas Health Care, Levine Childrens Hospital, Charlotte, North Carolina; <sup>20</sup> Centro de Desenvolvimento da Criança-Pediatric Hospital - CHUC, Coimbra, Portugal; <sup>21</sup> Division of Pediatric Neurology, Washington University, St. Louis, Missouri; <sup>22</sup> Department of Pediatrics, Section of Metabolic Disease, The Children's Hospital of Philadelphia, Perelman School of Medicine at the University of Pennsylvania Philadelphia, Pennsylvania; <sup>23</sup> Centre for Inherited Metabolic Diseases, Department of Clinical and Experimental Medicine, University of Catania, Catania, Italy; <sup>24</sup> NIH Undiagnosed Diseases Program, Common Fund, Office of the Director, NIH and National Human Genome Research Institute, NIH, Bethesda, Maryland; <sup>25</sup> Hunter Genetics, Waratah, New South Wales, School of Medicine and Public Health, University of Newcastle, Callaghan, New South Wales, Australia; <sup>26</sup> Department of Pathology and Laboratory Medicine, The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania; <sup>27</sup> Division of Genetics & Metabolism, Phoenix Children's Hospital, Phoenix, Arizona; <sup>28</sup> Pediatric Neurology, Stormont-Vail Health Care, Topeka, Kansas; <sup>29</sup> Department of Pediatrics, Odense University Hospital, Odense, Denmark; <sup>30</sup> Department of Neurology, Royal Children's Hospital, Parkville, Victoria, Australia; <sup>31</sup> Human Genetics and Genomic Medicine, University of Southampton and Wessex Clinical Genetics Service, Southampton, United Kingdom; <sup>32</sup> National Human Genome Research Institute, NIH, Bethesda, Maryland; <sup>33</sup> Department of Metabolic Medicine, Royal Children's Hospital, Murdoch Childrens Research Institute, Parkville, Victoria, Australia; <sup>34</sup> Department of Pediatric Neurology and Metabolism, University Hospital of Brussels, Brussels, Belgium; <sup>35</sup> Department of Molecular and Cellular Biochemistry, College of Medicine, University of Kentucky, Lexington, Kentucky; <sup>36</sup> Department of Genetics, University of Alabama at Birmingham, Birmingham, Alabama; <sup>37</sup> Howard Hughes Medical Institute, University of Washington, Seattle, Washington; <sup>38</sup> Carolinas Healthcare System, Charlotte, North Carolina; <sup>39</sup> Division of Genetics and Metabolism, University of South Florida, Tampa, Florida; <sup>40</sup> Department of Human Genetics, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, California; <sup>41</sup> Department of Pediatrics, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, California; <sup>42</sup> Department of Pediatrics, Hospital of Ostfold N-1603 Fredrikstad, Norway; <sup>43</sup> Department of Genetics and Genomic Sciences Icahn School of Medicine at Mount Sinai, New York, New York

Communicated by David S. Rosenblatt

Received 17 December 2015; accepted revised manuscript 17 February 2016.

Published online in Wiley Online Library (www.wiley.com/humanmutation). DOI: 10.1002/humu.22983

Additional Supporting Information may be found in the online version of this article.

\*Correspondence to: Hudson H. Freeze, Human Genetics Program, Sanford Children's Health Research Center, Sanford-Burnham-Prebys Medical Discovery Institute, 10901 N. Torrey Pines Rd. La Jolla, CA 92037. E-mail: hudson@sbpdiscovery.org

Contract grant sponsors: The Rocket Fund; National Institutes of Health (NIH) (R01DK095551, R01DK55615); National Human Genome Research Institute; National Heart Lung Blood Institute (1U54HG006493); NIGMS (GM102129) ERA-Net for Research Programs on Rare Diseases Joint Transnational Call 2011 (EURO-CDG) (ERARE11-135).

**ABSTRACT:** Congenital disorders of glycosylation (CDG) arise from pathogenic mutations in over 100 genes leading to impaired protein or lipid glycosylation. *ALG1* encodes a  $\beta$ 1,4 mannosyltransferase that catalyzes the addition of the first of nine mannose moieties to form a dolichol-lipid linked oligosaccharide intermediate required for proper

**N-linked glycosylation.** *ALG1* mutations cause a rare autosomal recessive disorder termed ALG1-CDG. To date 13 mutations in 18 patients from 14 families have been described with varying degrees of clinical severity. We identified and characterized 39 previously unreported cases of ALG1-CDG from 32 families and add 26 new mutations. Pathogenicity of each mutation was confirmed based on its inability to rescue impaired growth or hypoglycosylation of a standard biomarker in an *alg1*-deficient yeast strain. Using this approach we could not establish a rank order comparison of biomarker glycosylation and patient phenotype, but we identified mutations with a lethal outcome in the first two years of life. The recently identified protein-linked xeno-tetrasaccharide biomarker, NeuAc-Gal-GlcNAc<sub>2</sub>, was seen in all 27 patients tested. Our study triples the number of known patients and expands the molecular and clinical correlates of this disorder.

Hum Mutat 00:1–8, 2016. © 2016 Wiley Periodicals, Inc.

**KEY WORDS:** CDG; asparagine-linked glycosylation protein 1; carbohydrate-deficient transferrin; xeno-tetrasaccharide

## Introduction

Congenital disorders of glycosylation (CDG) are a group of predominantly autosomal recessive diseases characterized by broad multisystem defects; nearly all have neurological involvement [Freeze et al., 2012, 2015]. Disruption of the N-linked glycosylation pathway can cause more than 50 genetic disorders, but screening can be easily accomplished by analyzing the glycosylation status of the abundant serum glycoprotein, transferrin (Tf; MIM# 190000) [Freeze et al., 2014, 2015]. An abnormal serum sialotransferrin pattern provides clues as to which portion of the pathway is affected, but it typically cannot pinpoint the specific gene defect. Type I CDG's involve defects in either the synthesis or transfer of a dolichol lipid-linked oligosaccharide (DLO) [dolichol-P-P GlcNAc<sub>2</sub>Man<sub>9</sub>Glc<sub>3</sub> (2 N-acetylglucosamine (GlcNAc), 9 mannose (Man), and 3 glucose (Glc)] to nascent proteins in the endoplasmic reticulum (ER) [Freeze and Schachter, 2009]. An inability to efficiently synthesize or transfer full-sized DLO results in under-occupied glycosylation sites [Freeze and Schachter, 2009]. Defects in Golgi-dependent processing of the protein bound oligosaccharide or in Golgi-associated multifunctional protein complexes can result in type II CDG's [Freeze and Schachter, 2009]. Abnormal glycosylation of serum Tf is the most convenient indicator of deficient N-glycosylation in CDG patients, but some with confirmed CDG can have normal Tf [Freeze et al., 2012, 2–15].

By far, the most common CDG type involving the N-linked pathway is PMM2-CDG (MIM# 212065) with >800 (reported and unreported) affected individuals [Freeze et al., 2012]. However, in the vast majority of N-linked disorders, fewer than a dozen cases have been identified [Freeze et al., 2012].

Asparagine-linked glycosylation protein 1 homolog (*ALG1*; MIM# 605907) encodes an ER localized  $\beta$ 1,4 mannosyltransferase that catalyzes the transfer of the first of nine Man moieties onto the growing DLO [Huffaker and Robbins, 1982] (Fig. 1a). Mutations in *ALG1* cause a rare autosomal recessive disorder termed ALG1-CDG (formerly CDG-Ik; MIM# 608540) originally described by three independent groups who characterized four unrelated

individuals [Grubenmann et al., 2004; Kranz et al., 2004; Schwarz et al., 2004]. Since then, an additional 14 cases in 10 families have been described with a broad clinical spectrum ranging from mild intellectual disability to death within the first few weeks of life [Dupre et al., 2010; Morava et al., 2012; Snow et al., 2012; Rohlfing et al., 2014]. Here, we present an additional 39 previously unreported cases of ALG1-CDG from 32 families highlighting 26 new pathogenic mutations.

## Materials and Methods

### Subjects and Clinical Information

Families included in this research study provided written consent under an approved Sanford Burnham Prebys Medical Discovery Institute IRB protocol. Inclusion criteria for this study required at least one of the following: (1) the occurrence of at least one abnormal carbohydrate-deficient transferrin (CDT) test result indicating a type I CDG, or (2) molecular findings showing the presence of homozygous or compound heterozygous variants in *ALG1*. Several individuals appeared to have similar facial characteristics and consent was provided for including pictures from these and other affected individuals (Supp. Fig. S1).

### Molecular Analysis of *ALG1*

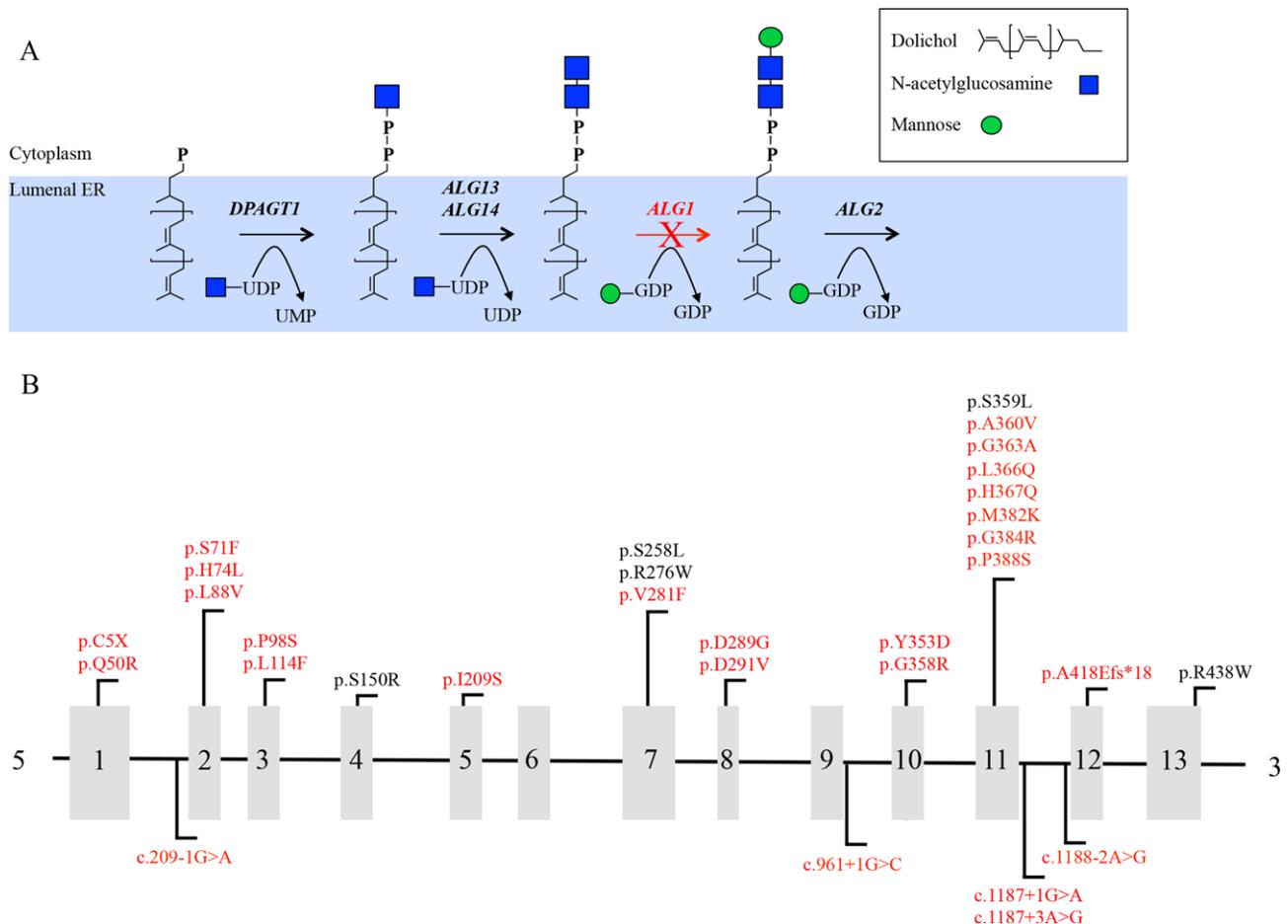
Mutation analysis for *ALG1* (NM\_019109.4, ENST00000262374) was performed at multiple centers by either direct gene sequencing of all 13 coding exons, including the exon–intron boundaries or in the instances where NGS panels or exome sequencing identified potential variants, only those exons were Sanger-confirmed. All primers were designed to exclude *ALG1* pseudogenes. All variants have been deposited in the LOVD3.0 ([https://grenada.lumc.nl/LOVD2/mendelian\\_genes/home.php?select\\_db=ALG1](https://grenada.lumc.nl/LOVD2/mendelian_genes/home.php?select_db=ALG1)).

### Yeast Complementation Assay

Human *ALG1* cDNA (RC206343 - Origene, Rockville, MD) was cloned into the yeast expression construct pYES2.1 (Thermo Fisher, Waltham, MA) to generate pYES2.1-*hALG1*. Before proceeding with complementation studies, all non-synonymous SNPs within the expression construct were removed so that only the human reference sequence (NM\_019109.4) was used. Subsequently, all patient-relevant missense mutations were generated using QuickChange Lightening (Agilent Technologies, Palo Alto, CA). Growth complementation and carboxypeptidase Y (CPY) glycosylation analysis of the *alg1*-deficient yeast strain was previously described [Grubenmann et al., 2004; Schwarz et al., 2004]. All yeast assays were replicated independently on three different occasions.

### Analysis of Xeno-Tetrasaccharide

As previously described [Bengtson et al., 2015; Zhang et al., 2015]. Biological samples tested for the xeno-tetrasaccharide included serum and/or plasma and primary fibroblast.



**Figure 1.** Pathway showing the enzymatic step requiring ALG1 mannosyltransferase activity and schematic showing mutations identified within *ALG1*. **A:** Schematic representation showing only the early N-linked glycosylation pathway on the cytoplasmic facing side of the ER with the ALG1-dependent step highlighted with a red (X) over the arrow. **B:** Schematic representation showing exon location of all the *ALG1* mutations identified in this study. The mutations identified in this study are highlighted in red. For splicing mutations, nucleotide numbering for cDNA uses +1 as the A of the ATG translation initiation codon in the reference sequence.

## Results

### Mutation Analysis

We utilized a combination of exome sequencing, targeted gene panels, and Sanger sequencing to identify a total of 31 potential mutations in 39 affected individuals of which 26 out of 31 (84%) were not previously reported to cause ALG1-CDG (Fig. 1b and Table 1). The five known mutations were p.Ser150Arg, p.Ser258Leu, p.Arg276Trp, p.Ser359Leu, and p.Arg438Trp. We searched the Exome Aggregation Consortium (ExAC) database (<http://exac.broadinstitute.org>) [v0.3 accessed 10.11.2015] of 60,706 unrelated individuals and found that nine out of 31 (35%) of these variants occurred at very low frequencies in heterozygous carriers, but 22 out of 31 (65%) had not been reported previously (Supp. Table S1). The most frequently observed mutation, in both the homozygous and compound heterozygous state (1 out of 39, 44%) (Table 1) was c.773C > T, which encodes a known pathogenic mutation p.Ser258Leu (rs28939378) and has been suggested to be a possible founder mutation [Dupre et al., 2010]. All six individuals homozygous for the p.Ser258Leu died before 6 months of age.

### Complementation of the $\Delta$ *alg1* Growth and CPY Glycosylation Defects by hALG1 Mutations

We determined the pathogenicity of unreported mutations using a temperature-sensitive (TS) *alg1*-deficient yeast strain [Grubenmann et al., 2004]. When grown at 26°C, this strain exhibits normal growth and appropriate glycosylation of an endogenous yeast glycoprotein, CPY [Grubenmann et al., 2004]. However, when grown at 37°C, this mutant expresses a p.Trp434X mutation leading to a C-terminal truncation of 16 amino acids. This results in growth suppression and CPY hypoglycosylation lacking between 1 and 4 of its N-glycan chains [Couto et al., 1984; Gao et al., 2004; Grubenmann et al., 2004]. In total, we tested 30 missense variants, including 11 reported disease-causing missense mutations and an additional 19 unreported variants. In silico pathogenicity prediction programs, Polyphen2 (<http://genetics.bwh.harvard.edu/pph2/>) and SIFT (<http://sift.jcvi.org>) gave mixed results for both novel and known mutations, despite nearly all our mutations occurring at highly conserved amino acids (Supp. Table S1). We also applied a third prediction method to score the impact of our variants: Combined Annotation Dependent Depletion (CADD) (<http://cadd.gs.washington.edu/>) integrates multiple annotations

**Table 1. Genotypes from 39 ALG1-CDG Cases**

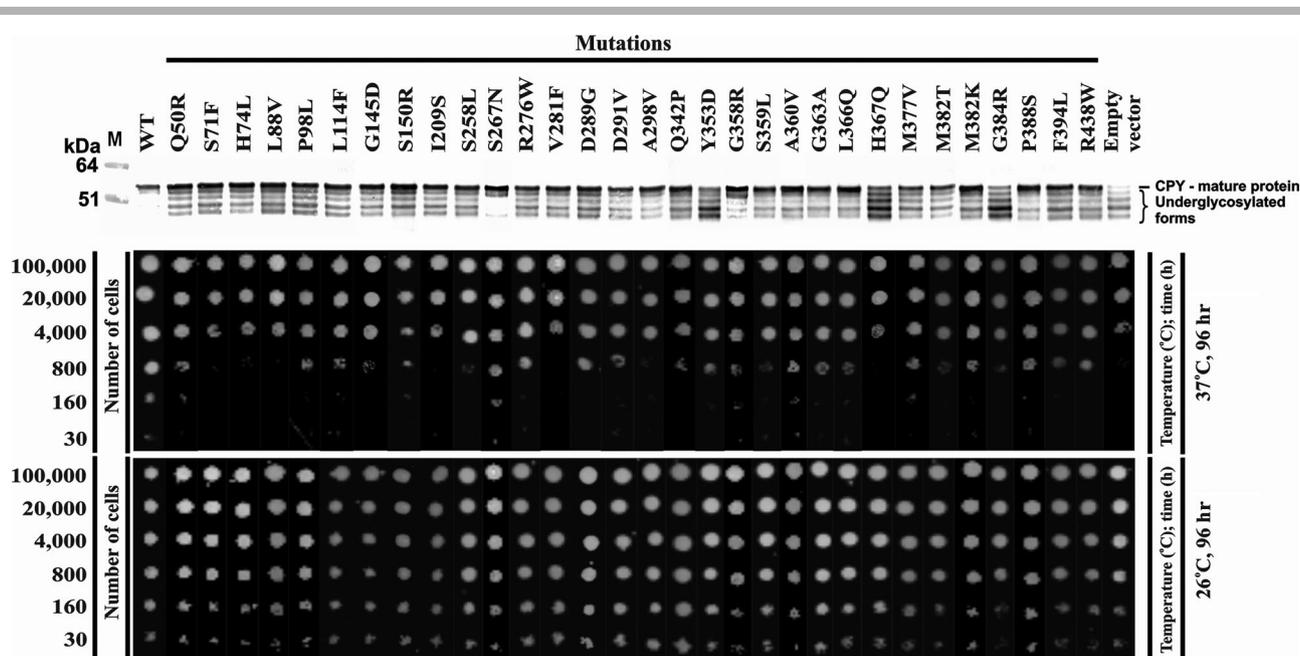
|            | Mutation 1                |                           | Mutation 2        |                              | Detection method |
|------------|---------------------------|---------------------------|-------------------|------------------------------|------------------|
|            | cDNA                      | Protein                   | cDNA              | Protein                      |                  |
| Patient 1  | c.866A > G                | p.Asp289Gly               | c.1312C > T       | p.Arg438Trp                  | Sanger           |
| Patient 2  | c.841G > T                | p.Val281Phe               | c.1101C > G       | p.His367Gln                  | Sanger           |
| Patient 3  | c.773C > T                | p.Ser258Leu               | c.773C > T        | p.Ser258Leu                  | Sanger           |
| Patient 4  | c.773C > T                | p.Ser258Leu               | c.773C > T        | p.Ser258Leu                  | Sanger           |
| Patient 5  | c.293C > T                | p.Pro98Leu                | c.773C > T        | p.Ser258Leu                  | Sanger           |
| Patient 6  | c.293C > T                | p.Pro98Leu                | c.773C > T        | p.Ser258Leu                  | Sanger           |
| Patient 7  | c.773C > T                | p.Ser258Leu               | c.773C > T        | p.Ser258Leu                  | Sanger           |
| Patient 8  | c.773C > T                | p.Ser258Leu               | c.1162C > T       | p.Pro388Ser                  | Sanger           |
| Patient 9  | c.841G > T                | p.Val281Phe               | c.1057T > G       | p.Tyr353Asp                  | WES              |
| Patient 10 | c.866A > G                | p.Asp289Gly               | c.866A > G        | p.Asp289Gly                  | Sanger           |
| Patient 11 | c.212C > T,<br>c.221A > T | p.Ser71Phe,<br>p.His74Leu | c.1079C > T       | p.Ala360Val                  | WES              |
| Patient 12 | c.149A > G                | p.Gln50Arg                | c.773C > T        | p.Ser258Leu                  | Sanger           |
| Patient 13 | c.149A > G                | p.Gln50Arg                | c.773C > T        | p.Ser258Leu                  | Sanger           |
| Patient 14 | c.773C > T                | p.Ser258Leu               | c.773C > T        | p.Ser258Leu                  | Sanger           |
| Patient 15 | c.1079C > T               | p.Ala360Val               | c.1187+1G > A     | NA                           | Sanger           |
| Patient 16 | c.1079C > T               | p.Ala360Val               | c.1187+1G > A     | NA                           | Sanger           |
| Patient 17 | c.149A > G                | p.Gln50Arg                | c.626T > G        | p.Ile209Ser                  | WES              |
| Patient 18 | c.773C > T                | p.Ser258Leu               | c.773C > T        | p.Ser258Leu                  | WES              |
| Patient 19 | c.342G > C                | p.Leu114Phe               | c.773C > T        | p.Ser258Leu                  | WES              |
| Patient 20 | c.262T > G                | p.Leu88Val                | c.1188-2A > G     | NA                           | Sanger           |
| Patient 21 | c.15C > A                 | p.Cys5Ter                 | c.149A > G        | p.Gln50Arg                   | Sanger           |
| Patient 22 | c.773C > T                | p.Ser258Leu               | c.1187+3A > G     | NA                           | Sanger           |
| Patient 23 | c.149A > G                | p.Gln50Arg                | c.1072G > C       | p.Gly358Arg                  | Sanger           |
| Patient 24 | c.773C > T                | p.Ser258Leu               | c.1079C > T       | p.Ala360Val                  | WES              |
| Patient 25 | c.773C > T                | p.Ser258Leu               | c.773C > T        | p.Ser258Leu                  | Capture          |
| Patient 26 | c.1076C > T               | p.Ser359Leu               | c.1250..1251insTG | p.Ala418Glu <sup>fs*18</sup> | Capture          |
| Patient 27 | c.872A > T                | p.Asp291Val               | c.872A > T        | p.Asp291Val                  | Capture          |
| Patient 28 | c.1076C > T               | p.Ser359Leu               | c.1076C > T       | p.Ser359Leu                  | Capture          |
| Patient 29 | c.626T > G                | p.Ile209Ser               | c.773C > T        | p.Ser258Leu                  | Capture          |
| Patient 30 | c.1097T > A               | p.Leu366Gln               | c.1097T > A       | p.Leu366Gln                  | Capture          |
| Patient 31 | c.1097T > A               | p.Leu366Gln               | c.1097T > A       | p.Leu366Gln                  | Sanger           |
| Patient 32 | c.450C > A                | p.Ser150Arg               | c.1312C > T       | p.Arg438Trp                  | WES              |
| Patient 33 | c.450C > A                | p.Ser150Arg               | c.1312C > T       | p.Arg438Trp                  | WES              |
| Patient 34 | c.209-1G > C              | NA                        | c.1079C > T       | p.Ala360Val                  | WES              |
| Patient 35 | c.209-1G > C              | NA                        | c.1079C > T       | p.Ala360Val                  | WES              |
| Patient 36 | c.1088G > C               | p.Gly363Ala               | c.1150G > A       | p.Gly384Arg                  | WES              |
| Patient 37 | c.773C > T                | p.Ser258Leu               | c.826C > T        | p.Arg276Trp                  | Sanger           |
| Patient 38 | c.773C > T                | p.Ser258Leu               | c.1145T > A       | p.Met382Lys                  | Sanger           |
| Patient 39 | c.961+1G > C              | NA                        | c.961+1G > C      | NA                           | Sanger           |

Genotypes identified in 39 ALG1-CDG cases. Only the variants believed to be pathogenic mutations are listed. However, in addition to the listed mutations, patients 11 and 20 were carriers of the p.Thr64Asn variant and patients 15 and 16 were carriers of a p.Pro346Arg variant. Detecting method is noted as well. Nucleotide numbering for cDNA uses +1 as the A of the ATG translation initiation codon in the reference sequence, with the initiation codon as codon 1. *ALG1* NCBI Accession (NM\_019109.4) and ENSEMBL (ENST00000262374).

into a single metric to predict and prioritize disease causing variants [Kircher et al., 2014]. These three programs were used as a guide for prioritizing which variants should be tested in our yeast model. However, since Polyphen2 and SIFT gave a number of conflicting results, we opted to test all the identified variants. Transformation of the mutant strain with each of the 30 individual missense variants followed by plating serial dilutions of the initial inoculum showed that none was able to fully rescue growth at the restrictive temperature when compared with wild-type ALG1, suggesting that each variant was, in fact, a pathogenic mutation (Fig. 2). At low dilutions on agar plates all mutants supported some growth, but differences were evident when 800 cells were spotted (Fig. 2). The visible growth suggests that either the mutant ALG1 protein retained residual activity or perhaps there was still endogenous yeast *alg1* residual activity remaining in a small subset of cells. However, at lower cell densities (160 and 30 cells) most showed no growth. A small set of mutations (H74L, G145D) were listed as benign or tolerated yet did not support growth, where as one mutant (L114F) was listed as probably damaging/damaging but only had a mild growth defect. Based on these

results, CADD was superior to Polyphen2 or SIFT in predicting the impact of ALG1 variants

CPY is used to monitor hypoglycosylation as it has four occupied glycosylation sites. Hypoglycosylation can result in the absence of 1–4 glycans and each glycan-deficient band is clearly visible in Western blots. *alg1*-deficient complemented with different hALG1 variants were initially grown at 26°C and transferred to 37°C. Before analysis by Western blotting, none of the mutants fully corrected CPY hypoglycosylation (Fig. 2), whereas wild-type hALG1 cDNA fully restored both normal growth and full CPY glycosylation (Fig. 2). We included a common polymorphism (p.Ser267Asn) (Fig. 2) to show our assay can differentiate pathogenic and benign variants. The hypoglycosylation patterns varied among the different mutants and we tried to rank order the degree of CPY hypoglycosylation using a weighted scale based on the proportion of molecules lacking the most N-glycans, for example, Y353D would be most severe, P388S least. Neither growth nor weighted CPY hypoglycosylation analysis allowed a rank order of which mutations are most damaging.



**Figure 2.** Yeast complementation assay using the  $\Delta alg1$  deficient yeast strain. The *alg1*-deficient yeast strain was transformed with the indicated pYES2.1 construct expressing either human wild-type *ALG1* or the indicated missense variants and allowed to grow for 96 hr at both permissive (26°C) and restrictive temperatures (37°C). Western blot analysis of carboxypeptidase Y (CPY) glycosylation was performed as previously described [Grubenmann et al., 2004]. Both CPY glycosylation and growth complementation were performed in triplicate with representative data presented.

There was also no correlation between patient clinical severity and the degree of CPY hypoglycosylation (data not shown). Furthermore, we know that in humans, homozygosity for p.Ser258Leu is always lethal within the first several months of life, yet it was not the most severe mutation in either yeast assays. For these reasons, we feel the yeast model should be used as a tool for determining pathogenicity, not clinical severity of potential mutants.

### Clinical Summary

In total, we studied 39 patients (17 male/22 female) from 32 unrelated families. Each of the 39 patients had at least one confirmed abnormal CDT test result. CDT is a general term used to describe a number of methods for analyzing serum Tf glycosylation status; methods used in this study included electrospray ionization mass spectrometry and isoelectric focusing.

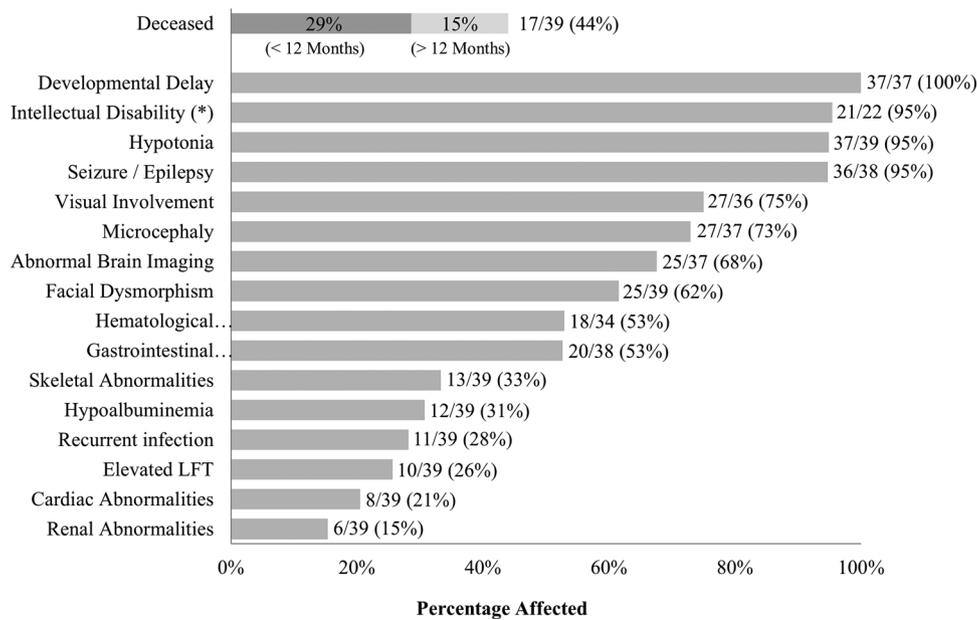
Patients had variable degree of neurodevelopmental deficiencies including developmental delay 37 out of 37 (100%), hypotonia 37 out of 39 (95%), seizures/epilepsy 36 out of 38 (95%), microcephaly 27 out of 37 (73%), abnormal brain imaging 25 out of 37 (68%) which consisted primarily of cerebral or cerebellar atrophy 11 out of 25 (44%), and for those who could be evaluated, intellectual disability 21 out of 22 (95%) (Fig. 3 and Supp. Table S2). These findings are consistent with previously reported cases. Ocular abnormalities that mainly involved strabismus 10 out of 27 (37%) and nystagmus six out of 27 (22%) were found in 27 out of 36 (75%) (Fig. 3 and Supp. Table S2). We also report a substantial number of patients with dysmorphic facial features 24 out of 39 (62%), hematological defects 18 out of 34 (53%), gastrointestinal problems 20 out of 38 (53%), skeletal abnormalities 13 out of 39 (33%), and hypoalbuminemia 12 out of 39 (31%) (Fig. 3 and Supp. Table S2). Hypoalbuminemia is noteworthy because all 12 of these individuals died at an average age of 6.75 months. Within this group, protein losing enteropathy (PLE) was documented in two individuals,

whereas three had renal disease. Furthermore, six out of 12 individuals presenting with hypoalbuminemia were homozygous for the p.Ser258Leu. Specifically, gastrointestinal manifestations were most often chronic diarrhea (seven out of 20) and/or PLE (five out of 20) (Supp. Table S2). Skeletal abnormalities consisted of scoliosis (five out of 13), kyphosis (two out of 13), or joint contractures (three out of 13) (Supp. Table S2). The phenotypes seen in ALG1-CDG individuals are not unique to this CDG type. Nearly all can be found in PMM2-CDG cases; whereas other characteristics such as PLE are more commonly seen in MPI-CDG, ALG6-CDG, and ALG8-CDG [Freeze et al., 2014].

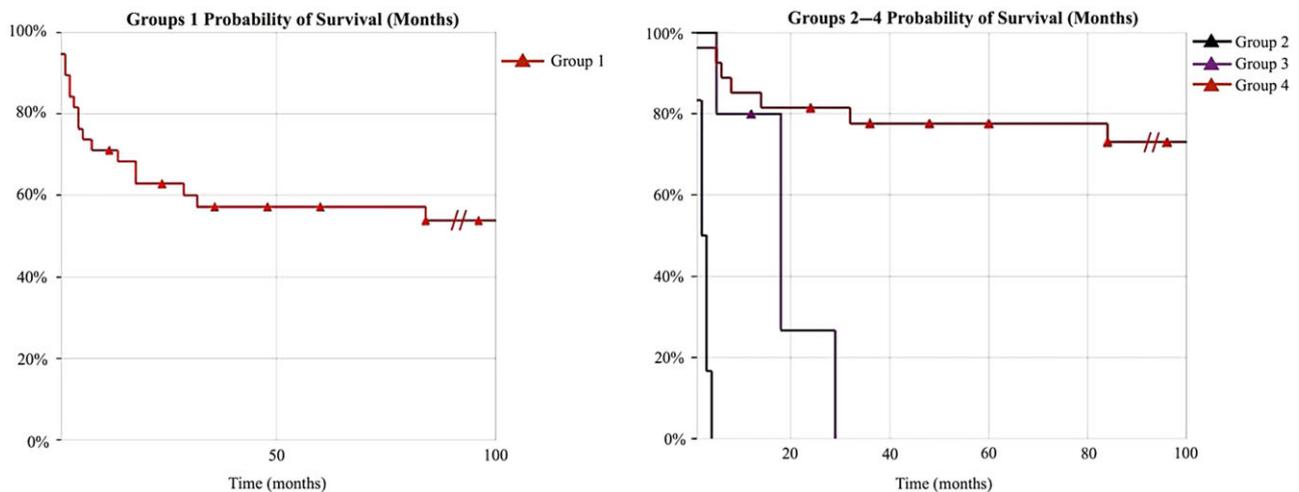
Premature death occurred in 17 out of 39 (44%) cases (Fig. 3) for a number of reasons including respiratory or renal failure and in several instances various infections leading to sepsis. The age distribution for those individuals who died ranged from <12 months of age in 11 out of 17 (65%) to >12 months of age in six out of 17 (35%) (Fig. 3 and Supp. Table S2). Kaplan–Meier estimator curves show that individuals homozygous for p.Ser258Leu, or compound heterozygous for the p.Gln50Arg are very likely to have lethal outcomes (Fig. 4). Premature death was previously reported in eight out of 18 (44%) carrying seven different mutations [Grubenmann et al., 2004; Kranz et al., 2004; Schwarz et al., 2004; Dupre et al., 2010; Morava et al., 2012; Snow et al., 2012; Rohlfing et al., 2014].

All six individuals homozygous for the p.Ser258Leu mutation died within the first 5 months of life. Two previously reported cases homozygous for this mutation died at 2 and 11 weeks of life [Kranz et al., 2004; Schwarz et al., 2004]. These individuals had characteristic dysmorphic facial features (Supp. Fig. S1). Four of the five individuals who were compound heterozygous for the p.Gln50Arg allele died between 5 and 28 months of age, irrespective of the second mutation. These four individuals also had similar facial features compared to other ALG1-CDG individuals with different genotypes (Supp. Fig. S1). One individual who is compound

## Clinical Summary of ALG1-CDG Cases



**Figure 3.** Phenotype summary from the 39 ALG1-CDG cases. Phenotypic data were provided for each individual and summarized as a percentage of patients affected. The number of deceased individuals is further broken down by the age at which the individual passed away. (\*) Indicates that not all individuals could be assessed for intellectual disability, due to early death or young age.



**Figure 4.** Kaplan–Meier survival curves for ALG1-CDG patients. **A:** Probability of survival for all 39 ALG1-CDG cases “Group 1”. **B:** Probability of survival for all 39 ALG1-CDG cases separated by genotype. Group 2 are individuals homozygous for the p.Ser258Leu ( $n = 6$ ), Group 3 are individuals compound heterozygous for the p.Gln50Arg ( $n = 5$ ), and Group 4 comprises the remaining individuals ( $n = 28$ ). In order to see the effect of the p.Ser258Leu and p.Gln50Arg, we made the time scale limit 100 months, however it should be noted that several individuals survived pass 100 months and this is denoted using (/).

heterozygous for the p.Gln50Arg remains alive at 9 months of age.

## Discussion

We report 39 individuals from 32 unrelated families with ALG1-CDG, more than tripling the number of known patients. To better characterize this disorder, we attempted to create an ALG1-deficient human cell line utilizing CRISPR/CAS9 technology to test

pathogenicity of the variants. However, we were unable to generate a cell line carrying a homozygous INDEL despite targeting four different exons (data not shown). This suggests that this gene is essential for viability.

Because we were unable to generate a null human cell line for complementation studies, we used a TS *alg1* mutant yeast strain. At the restrictive temperature, this mutant not only accumulates DLO containing dolichol-PP-GlcNAc<sub>2</sub>, but the GlcNAc<sub>2</sub> can be transferred to yeast glycoproteins such as exoglucanase [Cueva et al., 1998]. Various human cell lines grown under glucose-deprived conditions

or siRNA knockdown of *ALG1*, as well as the yeast *alg1* mutant, all accumulated proteins with GlcNAc<sub>2</sub> (Isono, 2011). This suggests that severely truncated oligosaccharides from ALG1-deficient cells can still be transferred to nascent proteins. It remains unclear if the majority of glycoproteins from ALG1-deficient cells bear this unusual glycan or if it only affects selected glycoproteins.

These observations in a model system have become more important because ALG1-CDG patients are known to accumulate a novel N-linked xeno-tetrasaccharide NeuAc-Gal-GlcNAc<sub>2</sub> [2 N-acetylglucosamine (GlcNAc), 1 galactose (Gal), 1 sialic acid (NeuAc)] on both total serum glycoproteins and on ~2%–8% of purified serum Tf [Bengtson et al., 2015; Zhang et al., 2015]. The presence of this tetrasaccharide shows that Dol-PP-GlcNAc<sub>2</sub> is flipped from the cytoplasmic face into the luminal face of the ER and subsequently transferred to proteins via the oligosaccharyltransferase complex. Once transferred, modified proteins enter the Golgi where GlcNAc<sub>2</sub> serves as an acceptor substrate for  $\beta$ 1,4 galactosyltransferase (MIM# 137060). Finally, Gal-GlcNAc<sub>2</sub> is capped with a NeuAc added by 2,6 sialyltransferase to form the final product NeuAc<sub>2</sub>,6-Gal  $\beta$ 1,4GlcNAc  $\beta$ 1,4GlcNAc $\beta$  [Bengtson et al., 2015; Zhang et al., 2015] (Supp. Fig. S2). Importantly, since this tetrasaccharide does not normally occur in mammals and is primarily detected in ALG1-CDG cases (trace amounts are detected in PMM2-CDG and MPI-CDG cases), it serves as a biomarker for either detecting or confirming a diagnosis of ALG1-CDG [Bengtson et al., 2015; Zhang et al., 2015]. We tested a limited number of type I CDG samples and cannot exclude the possibility that other type I's could have this tetrasaccharide. Yet, mechanistically it seems unlikely subsequent steps to ALG1 would have this specific glycan. In fact, of the 39 affected individuals enrolled into our study, 27 were tested and all had this novel tetrasaccharide present on either serum or fibroblast glycoproteins (Supp. Table S3). Of the 12 not tested, three were homozygous for the lethal p.Ser258Leu, four had previously reported mutations, and three were affected sibling pairs of tested patients. Two patients with novel mutations were not tested for the tetrasaccharide. We can confidently claim that this marker is associated with ALG1-CDG.

As seen with many glycosylation-related disorders, the clinical phenotype of ALG1-CDG varies dramatically between affected individuals (Fig. 3 and Supp. Table S2). While all ALG1-CDG patients do have substantial neurological involvement, other systems are variably affected (Fig. 3 and Supp. Table S2). This may be attributed to several factors including how ALG1 mutations affect residual mannosyltransferase activity or GDP-Man binding efficiency, complex formations with other proteins, tissue specific expression or even protein stability. For example, several ALG1 mutations including the p.Ser258Leu are highly unstable, resulting in near absence of protein expression in patient fibroblasts, whereas other mutations have no effect on stability (data not shown).

ALG1 is thought to form multiple protein complexes in yeast. Gao et al. (2004) show that ALG1 forms homo-oligomer structures and also interacts with subsequent  $\alpha$ -mannosyltransferases, ALG2 (MIM# 607905) and ALG11 (MIM# 613666). Together they add all four mannosyl-residues to the DLO on the cytoplasmic face of the ER [Gao et al., 2004]. With the discovery that mutations in *DPAGT1* (MIM#s 608093, 191350) and *ALG2* (MIM#s 607906, 616228) [Belaya et al., 2012; Cossins et al., 2013] can cause a CDG phenotype as well as a congenital myasthenic syndrome (CMS), it is conceivable that specific ALG1 mutations could also cause a mild CMS-like phenotype. Furthermore, traditional CDT testing could potentially miss such ALG1 cases given individuals with *DPAGT1*-CMS or *ALG2*-CMS had normal CDT testing [Belaya et al., 2012; Cossins et al., 2013].

We identified one case suspected of having ALG-CDG because of developmental delay, intellectual disability, hypotonia, epilepsy, and skeletal abnormalities. This individual was homozygous for a variant of unknown significance (VUS) (p.Thr64Asn) in *ALG1* which failed to rescue yeast growth or CPY glycosylation, suggesting it was deleterious. However, Tf glycosylation was normal and no tetrasaccharide was detected. Thus, she was excluded from our study. This finding shows the limitations of the yeast model and highlights the importance of having multiple methods to assess the pathogenicity of questionable variants.

In conclusion, we present molecular, clinical and biochemical findings in the largest collection of ALG1-CDG cases ever reported at a single time with 39 cases, bringing the total number of known ALG1-CDG to 57. This ranks it the third most common CDG type behind PMM2-CDG and ALG6-CDG [Freeze et al., 2012]. In addition, we identify highly lethal genotypes and confirm the presence of a unique xeno-tetrasaccharide in ALG1-CDG patients, thus showing it to be a biomarker for this disorder.

## Acknowledgments

We thank the families for their continued support. Dr. Markus Aebi for providing yeast strains and Dr. Claude Sansaricq for his contribution.

*Disclosure statement:* The authors declare no conflict of interest.

## References

- Belaya K, Finlayson S, Slater CR, Cossins J, Liu WW, Maxwell S, McGowan SJ, Maslau S, Twigg SR, Walls TJ, Pascual Pascual SI, Palace J, et al. 2012. Mutations in *DPAGT1* cause a limb-girdle congenital myasthenic syndrome with tubular aggregates. *Am J Hum Genet* 91:193–201.
- Bengtson P, Ng BG, Jaeken J, Matthijs G, Freeze HH, Eklund EA. 2015. Serum transferrin carrying the xeno-tetrasaccharide NeuAc-Gal-GlcNAc<sub>2</sub> is a biomarker of ALG1-CDG. *J Inher Metab Dis* 39:107–114.
- Cossins J, Belaya K, Hicks D, Salih MA, Finlayson S, Carboni N, Liu WW, Maxwell S, Zoltowska K, Farsani GT, Laval S, Seidhamed MZ, et al. 2013. Congenital myasthenic syndromes due to mutations in *ALG2* and *ALG14*. *Brain* 136:944–956.
- Couto JR, Huffaker TC, Robbins PW. 1984. Cloning and expression in *Escherichia coli* of a yeast mannosyltransferase from the asparagine-linked glycosylation pathway. *J Biol Chem* 259:378–382.
- Cueva R, Cotano C, Larrriba, G. 1998. N-glycosylation by transfer of GlcNAc<sub>2</sub> from dolichol-PP-GlcNAc<sub>2</sub> to the protein moiety of the major yeast exoglucanase. *Yeast* 14:773–781.
- Dupre T, Vuillaumier-Barrot S, Chantret I, Sadou Yaye H, Le Bizet C, Afenjar A, Altuzarra C, Barnerias C, Burglen L, de Lonlay P, Feillet F, Napuri S, et al. 2010. Guanosine diphosphate-mannose: GlcNAc<sub>2</sub>-PP-dolichol mannosyltransferase deficiency (congenital disorders of glycosylation type Ik): five new patients and seven novel mutations. *J Med Genet* 47:729–735.
- Freeze HH and Schachter H. 2009. Genetic disorders of glycosylation. In Varki A, Cummings RD, Esko JD, Freeze HH, Stanley P, Bertozzi CR, Hart GW, and Etzler ME, editors. *Essentials of glycobiology*. New York: Cold Spring Harbor. Chapter 42.
- Freeze HH, Eklund EA, Ng BG, Patterson MC. 2012. Neurology of inherited glycosylation disorders. *Lancet Neurol* 11:453–466.
- Freeze HH, Chong JX, Bamshad MJ, Ng BG. 2014. Solving glycosylation disorders: fundamental approaches reveal complicated pathways. *Am J Hum Genet* 94:161–175.
- Freeze HH, Eklund EA, Ng BG, Patterson MC. 2015. Neurological aspects of human glycosylation disorders. *Annu Rev Neurosci* 38:105–125.
- Gao XD, Nishikawa A, Dean N. 2004. Physical interactions between the *Alg1*, *Alg2*, and *Alg11* mannosyltransferases of the endoplasmic reticulum. *Glycobiology* 14:559–570.
- Grubenmann CE, Frank CG, Hulsmeier AJ, Schollen E, Matthijs G, Mayatepek E, Berger EG, Aebi M, Hennet T. 2004. Deficiency of the first mannosylation step in the N-glycosylation pathway causes congenital disorder of glycosylation type Ik. *Hum Mol Genet* 13:535–542.
- Huffaker TC, Robbins PW. 1982. Temperature-sensitive yeast mutants deficient in asparagine-linked glycosylation. *J Biol Chem* 257:3203–3210.

- Isono T. 2011. O-GlcNAc-specific antibody CTD110.6 cross-reacts with N-GlcNAc2-modified proteins induced under glucose deprivation. *PLoS One* 6: e18959.
- Kranz C, Denecke J, Lehle L, Sohlbach K, Jeske S, Meinhardt F, Rossi R, Gudowius S, Marquardt T. 2004. Congenital disorder of glycosylation type Ik (CDG-Ik): a defect of mannosyltransferase I. *Am J Hum Genet* 74:545–551.
- Kircher M, Witten DM, Jain P, O’Roak BJ, Cooper GM, Shendure J. 2014. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet* 46:310–315.
- Morava E, Vodopiutz J, Lefeber DJ, Janecke AR, Schmidt WM, Lechner S, Item CB, Sykut-Cegielska J, Adamowicz M, Wierzbza J, Zhang ZH, Mihalek I, et al. 2012. Defining the phenotype in congenital disorder of glycosylation due to ALG1 mutations. *Pediatrics* 130:e1034–1039.
- Rohlfing AK, Rust S, Reunert J, Tirre M, Du Chesne I, Wemhoff S, Meinhardt F, Hartmann H, Das AM, Marquardt T. 2014. ALG1-CDG: a new case with early fatal outcome. *Gene* 534:345–351.
- Schwarz M, Thiel C, Lubbehusen J, Dorland B, de Koning T, von Figura K, Lehle L, Korrner C. 2004. Deficiency of GDP-Man:GlcNAc2-PP-dolichol mannosyltransferase causes congenital disorder of glycosylation type Ik. *Am J Hum Genet* 74:472–481.
- Snow TM, Woods CW, Woods AG. 2012. Congenital disorder of glycosylation: a case presentation. *Adv Neonatal Care* 12:96–100.
- Zhang W, James PM, Ng BG, Li X, Xia B, Rong J, Asif G, Raymond K, Jones MA, Hegde M, Ju T, Cummings RD, et al. 2015. A novel N-tetrasaccharide in patients with congenital disorders of glycosylation including asparagine-linked glycosylation protein 1, phosphomannomutase 2, and phosphomannose isomerase deficiencies. *Clin Chem* 62:208–217.