Genotype-phenotype association and variant characterization in Diamond-Blackfan anemia caused by pathogenic variants in

M. D. Gianferante
M. W. Wlodarski
E. Atsidaftos
L. Da Costa
P. Delaporta

See next page for additional authors

Follow this and additional works at: https://academicworks.medicine.hofstra.edu/articles

Part of the Pediatrics Commons

Recommended Citation
Authors
Genotype-phenotype association and variant characterization in Diamond Blackfan anemia caused by pathogenic variants in RPL35A

D. Matthew Gianferante,1 Marcin W. Wlodarski,2,3 Evangelia Atsidaftos,4 Lydie Da Costa,2,4 Polyxeni Delaporta,1 Jason E. Farrar,3 Frederick D. Goldman,9 Maryam Hussain,1 Antonis Kattamis,1 Thierry Leblanc,4 Jeffrey M. Lipton,4 Charlotte M. Niemeyer,2 Dagmar Pospisilova,10 Paola Quarello,11 Ugo Ramenghi,12 Vijay G. Sankaran,13 Adrianna Vlachos,9 Jana Volejnikova,10 Blanche P. Alter,1 Sharon A. Savage1 and Neelam Giri1

1Clinical Genetics Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, NIH, Rockville, MD, USA; 2Division of Pediatric Hematology and Oncology, Department of Pediatrics and Adolescent Medicine, Medical Center, Faculty of Medicine, University of Freiburg, Freiburg, Germany; 3St. Jude Children’s Research Hospital, Memphis, TN, USA; 4Feinstein Institutes for Medical Research, Division of Hematology/Oncology and Cellular Therapy, Cohen Children’s Medical Center, Northwell Health, New York, NY, USA; 5Service d’Hématologie Biologique, Hôpital Robert-Debré, Université de Paris, Paris, France; 6Laboratory of Excellence for Red Blood Cells, GR-Ex, Paris, France; 7First Department of Pediatrics, National and Kapodistrian University of Athens, Athens, Greece; 8Arkansas Children’s Research Institute and University of Arkansas for Medical Sciences, Little Rock, AR, USA; 9University of Alabama at Birmingham, Birmingham, AL, USA; 10Palacky University and University Hospital, Olomouc, Czech Republic; 11Regina Margherita Children’s Hospital, Torino, Italy; 12Pediatric and Public Health Science, University of Torino, Torino, Italy and 13Division of Hematology/Oncology, Boston Children’s Hospital and Department of Pediatric Oncology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA, USA

ABSTRACT

Diamond Blackfan anemia (DBA) is predominantly an autosomal dominant inherited red cell aplasia primarily caused by pathogenic germline variants in ribosomal protein genes. DBA due to pathogenic RPL35A variants has been associated with large 3q29 deletions and phenotypes not common in DBA. We conducted a multi-institutional genotype-phenotype study of 45 patients with DBA associated with pathogenic RPL35A germline variants and curated the variant data on 21 additional cases from the literature. Genotype-phenotype analyses were conducted comparing patients with large deletions versus all other pathogenic variants in RPL35A. Twenty-two of the 45 cases had large deletions in RPL35A. After adjusting for multiple tests, a statistically significant association was observed between patients with a large deletion and steroid-resistant anemia, neutropenia, craniofacial abnormalities, chronic gastrointestinal problems, and intellectual disabilities (P<0.01) compared with all other pathogenic variants. Non-large deletion pathogenic variants were spread across RPL35A with no apparent hot spot and 56% of the individual family variants were observed more than once. In this, the largest known study of DBA patients with pathogenic RPL35A variants, we determined that patients with large deletions have a more severe phenotype that is clinically different from those with non-large deletion variants. Genes of interest also deleted in the 3q29 region that could be associated with some of these phenotypic features include LMLN and IQCG. Management of DBA due to large RPL35A deletions may be challenging due to complex problems and require comprehensive assessments by multiple specialists including immunological, gastrointestinal, and developmental evaluations to provide optimal multidisciplinary care.
Introduction

Diamond Blackfan anemia (DBA) is a rare inherited bone marrow (BM) failure syndrome (IMFS) with an estimated incidence of 5-10 per million live births.\textsuperscript{1,5} It is characterized by failure of red blood cell (RBC) production, congenital malformations, and cancer predisposition. Classic DBA consists of profound anemia diagnosed before one year of age, macrocytosis, reticulocytopenia, and a paucity of erythroid precursors in the BM.\textsuperscript{6} However, DBA is a heterogeneous disorder with many cases having additional symptoms or no symptoms at all. Congenital abnormalities are variable but include mostly midline craniofacial defects, renal, cardiac, and thumb abnormalities.\textsuperscript{1} Malignancies associated with DBA include myeloid neoplasia, colorectal adenocarcinoma, osteogenic sarcoma, and genitourinary cancers.\textsuperscript{1,9} Standard treatment for anemia in DBA is long-term corticosteroids. Nearly 40% of people with DBA who respond to initial treatment become steroid-dependent, and those who fail to respond to corticosteroids require chronic RBC transfusions and iron chelation or hematopoietic cell transplant (HCT).\textsuperscript{1,9} DBA is predominately an autosomal dominant disorder caused by pathogenic germline variants in genes encoding ribosomal proteins.\textsuperscript{10} Twenty-six ribosomal genes have been linked to DBA etiology as well as two X-linked genes, TSρ2 and CATA, which encode a ribosome chaperone and a hematopoietic transcription factor targeted by altered ribosome levels, respectively.\textsuperscript{10,11} Although all known DBA genes are typically included in gene discovery analysis or descriptive cohort studies of DBA, genotype-phenotype studies aimed to find clinical association by gene are currently limited. Only the most frequent ribosomal genes have been examined from a genotype-phenotype perspective (e.g., RPS19, RPL11, RPL5) with little to no information available for the majority of disease-causing genes, including RPL35A.\textsuperscript{10,12} RPL35A was first associated with DBA in 2008 and has been reported to cause about 3.5% of DBA cases.\textsuperscript{13,14} Typically, data on DBA due to RPL35A have been confined to case reports, or as part of a larger DBA study with limited to no phenotypic information.\textsuperscript{10,13,14}

RPL35A codes for a large ribosomal subunit protein located at the telomeric end of chromosome 3q (3q29-qter); pathogenic germline variants have been reported as single-nucleotide variants (SNV), small frameshifts, inframe deletions, and large deletions involving the entire RPL35A gene with or without multiple contiguous genes deleted in the 3q29 region.\textsuperscript{14,15} Patients with DBA caused by RPL35A have been reported to have severe anemia as well as additional phenotypes not usually shared by other DBA patients, such as immunodeficiency and autism spectrum disorder.\textsuperscript{12,13,15} It is unclear whether these phenotypes are associated with the deletion of RPL35A itself or other genes deleted in the region. Some of these phenotypes overlap with 3q29 deletion syndrome, a clinical syndrome in which the contiguous deleted region is near, but does not include, RPL35A.\textsuperscript{16} The phenotype of 3q29 deletion syndrome is thought to be related to the genes within the deleted region and includes dysmorphic facial features, intellectual disability, musculoskeletal problems, and neuropsychiatric issues.\textsuperscript{13,16,21} The deletion in 3q29 deletion syndrome is typically about 1.5 megabase (Mb) in size, and the consistent size is thought to be related to low copy repeat regions at each end of the deletion.\textsuperscript{12,13,21,24}

We conducted a multi-institutional international collaborative study of patients with DBA due to RPL35A to determine the clinical consequences of germline large deletions versus other pathogenic RPL35A variants. We also assembled variant data from additional published cases of RPL35A-associated DBA to better characterize pathogenic germline variants in this disease. The phenotypes of these patients were compared with those of the 3q29 deletion syndrome to elucidate the similarities and the differences related to variants in this region of the genome.

Methods

Study population

Patients with DBA due to RPL35A were identified within the National Cancer Institute (NCI) IMFS cohort (clinicaltrials.gov identifier: NCT00027274),\textsuperscript{20} and through collaboration with investigators from the DBA Registry of North America (DBAR. clinicaltrials.gov identifier: NCT00106015),\textsuperscript{20} DBA registries from Germany, France, Italy, the Czech Republic, and Greece, and through investigators from Alabama Children’s Hospital, Arkansas Children’s Hospital, and Boston Children’s Hospital (Figure 1 and Online Supplementary Table S1). All individuals were participants in Institutional Review Board approved protocols and had signed informed consents for participating in research studies.

Additional cases of DBA for RPL35A pathogenic variant characterization were identified through a search of PubMed and review of ClinVar\textsuperscript{23} and Human Gene Mutation Database (HGMD).\textsuperscript{28} All data were extracted as of February 22, 2019. ClinVar variants were restricted to pathogenic or likely pathogenic RPL35A DBA variants that met the minimum requirements for data sharing and quality assurance,\textsuperscript{27} and HGMD was restricted to RPL35A DBA disease-causing mutation. Any case that was a potential duplicate of a collaborator case was excluded (Figure 1).

Clinical data extraction

Data extraction focused on clinical criteria associated with DBA and with 3q29 deletion syndrome.\textsuperscript{22,23} A positive finding was counted as present, and a clinical finding marked absent or not stated was considered absent. Standard criteria for defining cytopenia and immunodeficiency were used, and definitions of phenotypes studied are outlined in Online Supplementary Table S2.\textsuperscript{22,23}

Pathogenic variant calling and variant annotation

Pathogenic variant locations (i.e., genomic or chromosomal coordinates) were used as provided by collaborators or from publications. Methods used to identify pathogenic variants included targeted sequencing, panel testing, exome sequencing, multiplex ligation-dependent probe amplification, SNP array, or array comparative genomic hybridization (Online Supplementary Table S3). Inclusion of a variant in the study required it to be reported as pathogenic, be rare (minor allele frequency [MAF] <1% within any gnomAD ethnic subgroups\textsuperscript{32}), missense variants needed to be predicted pathogenic by meta in silico predictor programs (CADD >25)\textsuperscript{33} and REVEL (>0.5)\textsuperscript{34} (Online Supplementary Table S3), and the same case could not be included more than once. Any variants of unclear pathogenicity, including untranslated regions (UTR) and duplication, were excluded. A “large deletion” was defined as a deletion of the entire RPL35A gene. “Small frameshift” included all insertion and deletions that were not large or inframe. SnpEff was used to annotate missing chromosomal, genomic, or protein positions.\textsuperscript{26} ANNOVAR was used to annotate MAF from publicly
available databases and obtain in silico prediction scores. The St. Jude Cloud PeCan tool was used to create lollipop plots. For large deletions, UCSC Table Browser was used to identify all deleted genes in that region.  

Genotype-phenotype statistical analysis

The cases for the genotype-phenotype analysis were made up exclusively from the ten collaborating institutions (Figure 1). The cases identified through literature review had limited to no available clinical data and were included in RPL35A variant characterization. All statistical analyses used the individual case as the analysis unit. We compared 21 different clinical features between cases with large deletions versus all other pathogenic variants using Fisher exact test or Student t-test (Stata 15 software). Bonferroni correction for 21 tests was applied. \( P<2.4\times10^{-3} \) were considered significant.

Results

Cohort clinical characterization and genotype-phenotype relationship

We compiled a total of 45 cases of DBA caused by RPL35A variants from ten different institutions/registries after exclusion of four cases with variants of unknown significance (Figure 1). Table 1 shows the clinical features of the 45 patients from 41 different families. Fifteen of these cases have been published previously (Online Supplementary Table S3). Nearly half of the patients (\( n=22 \)) had large deletions involving the entire RPL35A along with additional genes deleted in the region. All other pathogenic variants (\( n=23 \)) were either nonsense, missense, splice site, small frameshift, or intron deletions. Six of these variants are not present in ClinVar, HGMD, or gnomAD, and are considered novel (Online Supplementary Table S3).

Forty-four percent of the cohort was male (male:female 1:1.25; \( P=0.14 \)) (Table 1). Median age at identification of symptoms was at birth in patients with large deletions versus two months of age in those with other pathogenic variants (\( P=0.04 \)) (Table 1). Age of last follow-up was similar for both groups with a median age of 10 and 12 years, respectively (\( P=0.37 \) [data not shown]). All except two patients were alive at last follow-up. The cause of death was right ventricular outflow obstruction due to thromboembolic plaque in one patient and complications related to treatment for Wilms tumor in the other patient. Erythrocyte adenosine deaminase (eADA) values were available only for a subset of patients. Six of 11 patients with large deletion and 5 of 13 with all other RPL35A variants had elevated eADA with no difference between the groups (\( P=0.682 \)). Anemia was the most common presenting symptom and was identified in all except one patient. More than 86% of patients in both groups had severe anemia needing treatment (Table 1), but a significant number of patients with RPL35A large deletions were steroid-resistant and RBC transfusion-dependent compared with patients with other pathogenic variants in RPL35A (\( P=0.0023 \)) (Table 1). The number of patients who received HCT was similar in both groups. Neutropenia was significantly associated with large deletions (\( P=0.0017 \)). In addition, a higher proportion of patients with large deletions with severe neutropenia
includes missense, nonsense, small frameshift, splice, and intramere deletions. A positive finding was counted as present, and a clinical finding marked absent or not stated was analyzed.

### Hematologic and immune phenotypes

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>Cases with RPL35A large deletion</th>
<th>Cases with all other RPL35A variants</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anemia</td>
<td>44</td>
<td>22</td>
<td>1</td>
</tr>
<tr>
<td>Requiring treatment</td>
<td>42</td>
<td>22</td>
<td>0.23</td>
</tr>
<tr>
<td>HCT</td>
<td>7</td>
<td>4</td>
<td>0.70</td>
</tr>
<tr>
<td>Steroid-resistant</td>
<td>26</td>
<td>18</td>
<td>0.0023</td>
</tr>
<tr>
<td>Thrombocytopenia</td>
<td>4</td>
<td>1</td>
<td>0.61</td>
</tr>
<tr>
<td>Neutropenia</td>
<td>33</td>
<td>21</td>
<td>0.0017</td>
</tr>
<tr>
<td>Requiring treatment</td>
<td>10</td>
<td>9</td>
<td>0.004</td>
</tr>
<tr>
<td>Immunodeficiency</td>
<td>9</td>
<td>8</td>
<td>0.01</td>
</tr>
<tr>
<td>Recurrent infection</td>
<td>15</td>
<td>11</td>
<td>0.03</td>
</tr>
</tbody>
</table>

### Congenital abnormalities

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>Cases</th>
<th>Cases with RPL35A large deletion</th>
<th>Cases with all other RPL35A variants</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Craniofacial</td>
<td>12</td>
<td>11</td>
<td>1</td>
<td>0.0006</td>
</tr>
<tr>
<td>Microcephaly</td>
<td>10</td>
<td>7</td>
<td>3</td>
<td>0.17</td>
</tr>
<tr>
<td>Skeletal/limb</td>
<td>11</td>
<td>9</td>
<td>2</td>
<td>0.02</td>
</tr>
<tr>
<td>Cardiac</td>
<td>10</td>
<td>7</td>
<td>3</td>
<td>0.17</td>
</tr>
<tr>
<td>Urogenital</td>
<td>10</td>
<td>5</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>Short stature</td>
<td>23</td>
<td>11</td>
<td>12</td>
<td>1</td>
</tr>
<tr>
<td>Any congenital abnormality</td>
<td>26</td>
<td>16</td>
<td>10</td>
<td>0.07</td>
</tr>
<tr>
<td>≥3 abnormalities</td>
<td>13</td>
<td>11</td>
<td>2</td>
<td>0.003</td>
</tr>
</tbody>
</table>

### Other phenotypes

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>Cases</th>
<th>Cases with RPL35A large deletion</th>
<th>Cases with all other RPL35A variants</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chronic GI problem</td>
<td>9</td>
<td>9</td>
<td>0</td>
<td>0.0006</td>
</tr>
<tr>
<td>Intellectual disability</td>
<td>15</td>
<td>13</td>
<td>2</td>
<td>0.0004</td>
</tr>
</tbody>
</table>

### Notes

1. Absolute neutrophil count <0.5x10⁹/L required treatment for neutropenia compared with patients with other pathogenic variants (n=9 vs. n=1; P=0.004).
2. Treatment of neutropenia included cases with documented use of granulocyte-colony stimulating factor (GCSF) and treatment not specified in two cases. Bone marrow findings were available only in nine patients with large deletions and three with other pathogenic variants. Patients with large deletion and severe neutropenia (n=7) appeared to have both myeloid and erythroid hypoplasia while those with no neutropenia had mainly erythroid hypoplasia (5 of 7 vs. 0 of 3; P=0.06).

### Treatment and Follow-up

There were six cases with loss of function (LOF) variants (splice site=3, nonsense=1, frameshift=2) among patients with other pathogenic variants in RPL35A. Univariate analysis comparing these six cases with the 19 cases with missense or frameshift deletions showed an earlier age for the diagnosis of anemia in cases with LOF variants (P=0.02) but no association with neutropenia or immunodeficiency. Only one patient (frameshift variant) had severe neutropenia that was managed with GCSF, and one (splice site defect) had neutropenia and immunodeficiency treated with immunoglobulins.

Fifty-eight percent (n=26) of the cohort had at least one congenital abnormality involving craniofacial, skeletal and/or limb, cardiac, renal, genital abnormalities, or microcephaly (Table 1). The abnormalities described in each individual case are listed in Online Supplementary Table S4. Craniofacial abnormalities were significantly more frequent in cases with RPL35A large deletions compared with other pathogenic variants (n=11 vs. n=1; P=0.0006). Cases with large deletions were also more likely to have skeletal/limb abnormalities (n=9 vs. n=2; P=0.02) as well as three or more abnormalities (n=11 vs. n=2; P=0.008). Short stature, a common feature of DBA, was identified in approximately 50% of cases in each group and there was no difference in the use of steroids in...
variants are represented once per family. cDNA position based on transcript NM_000996.3. Numbers and letters indicate amino acids. Dotted lines indicate exons. Gray circles: literature, HGMD, and ClinVar cases. Black circles: collaborator cases. Cluster, defined as >5 unrelated cases by family in one area, is located between codon 28 to 33 (*). The numbers within the circles indicate the number of unique families with that specific pathogenic variant.

Figure 2. Distribution of RPL35A pathogenic variants reported in patients with Diamond-Blackfan anemia (DBA). All pathogenic single nucleotide and small indel variants are represented once per family. cDNA position based on transcript NM_000996.3. Numbers and letters indicate amino acids. Dotted lines indicate exons. Gray circles: literature, HGMD, and ClinVar cases. Black circles: collaborator cases. Cluster, defined as >5 unrelated cases by family in one area, is located between codon 28 to 33 (*). The numbers within the circles indicate the number of unique families with that specific pathogenic variant.

**Table 2. Pathogenic germ line RPL35A variant type.**

<table>
<thead>
<tr>
<th>Variant type</th>
<th>N.</th>
<th>Percent of families</th>
</tr>
</thead>
<tbody>
<tr>
<td>Large deletion</td>
<td>28</td>
<td>45%</td>
</tr>
<tr>
<td>Inframe deletion</td>
<td>10</td>
<td>16%</td>
</tr>
<tr>
<td>Missense</td>
<td>14</td>
<td>23%</td>
</tr>
<tr>
<td>Small frameshift</td>
<td>4</td>
<td>6%</td>
</tr>
<tr>
<td>Splice</td>
<td>4</td>
<td>6%</td>
</tr>
<tr>
<td>Nonsense</td>
<td>2</td>
<td>3%</td>
</tr>
</tbody>
</table>

Pathogenic variants were only counted once per family. Percent is based upon a total of 62 unique unrelated families in the study. Pathogenic variants must have been reported by collaborators as pathogenic and have minor allele frequency <1% in gnomAD. Variants of unknown significance including two UTR and two duplication variants were excluded. Large deletion was defined as entire RPL35A deleted, small frameshift included insertion/deletion that was not large or inframe, and no large insertions were identified.

either group. Delayed development and intellectual disability including learning problems requiring an individualized education program in school were significant among cases with RPL35A large deletions compared with other pathogenic variants (n=13 vs. n=2; P=0.0004) (Table 1); one case also had a diagnosis of autism spectrum disorder. Chronic gastrointestinal (GI) problems, including chronic diarrhea, gastrostomy feeding tube-dependence, and chronic enterocolitis, were also significantly associated with large deletions compared with all other pathogenic variants (n=9 vs. n=0; P=0.0006) (Table 1).

In an attempt to evaluate the phenotypes of cases with non-large deletion pathogenic variants based on their genomic location, we noted that the variants were spread across the RPL35A gene with one area of clustering from codon 28 to 33 (Figure 2). A cluster was defined as >5 unrelated cases in one area. There were nine cases from seven different families from collaborators in this cluster. The variants in six of the seven families were the same inframe deletion (c.82-84delCTT NM_000996.3). We did not find any difference in the hematologic or physical phenotype between the cases in the cluster and those spread across the RPL35A gene (data not shown).

**RPL35A pathogenic variant characterization**

In addition to the 45 cases obtained from collaboration with DBA registries/investigators used in the above genotype/phenotype analysis, we identified 21 other cases of DBA due to RPL35A pathogenic variants reported in the literature. We characterized the pathogenic variants in all 66 cases from 62 unrelated families. Two families had three affected family members each sharing a variant c.505 G>C or c.82-84CTT (NM_000996.3), respectively.

Twenty-eight of the 62 pathogenic RPL35A variants from unrelated families were large deletions (45%) (Table 2). Genomic co-ordinates were not available for six cases with large deletions. The size of the large deletions in 22 cases with available data ranged from 5kb to 13.42 Mb (Online Supplementary Table S5). The number of genes deleted in these cases ranged from two to 153. Large deletions extending upstream and downstream of the 5q29 deletion syndrome region were observed in 13 cases, four cases included part but not all of the region, and five did not include the region at all (Figure 3 and Online Supplementary Table S5).

Thirty-four of the 62 pathogenic variants were non-large deletions from 19 unique collaborator families and 15 unique families from literature cases. Fifty-six percent (19 of 34) of the unique family variants were observed more than once (Online Supplementary Table S6). The inframe deletion c.82-84delCTT was the most common pathogenic variant (n=10) (Online Supplementary Table S6), followed by missense c.125A>G (n=3); three other variants were reported in two families each. All other variants were spread across the gene without clear mutational hotspots and no additional clusters were identified using the pathogenic variants from the cases in the literature (Figure 2). Most of the variants were located in the middle and 3' end of the gene with only one pathogenic variant located near the beginning of the gene.

**Discussion**

We have compiled the largest cohort of patients with DBA due to pathogenic germ line RPL35A variants. This international multi-institutional collaboration allowed us to combine well-curated genomic information with extensive clinical data to better characterize the genotype and DBA phenotype associated with variants involving RPL35A. We conclude that cases with RPL35A large deletions are phenotypically different from cases with other variant types in RPL35A and have a more complex clinical phenotype.

Large deletions in RPL35A were frequent, comprising 45% of the pathogenic variants in this cohort. The fre-
frequency of large deletions in published cases of DBA caused by pathogenic variants in \textit{RPL35A} is unknown due to the rarity of the disease. However, our findings are consistent with what is reported in HGMD where 40% (8 of 21) of DBA cases caused by \textit{RPL35A} are reported as large/gross deletions. Fifty-six percent of non-large deletion pathogenic variants were comprised of five variants. Domain information is not well described for this gene, but we did identify apparent clustering of cases between codon 28 to 33 that may be indicative of a potentially important domain (Figure 2). A sufficient number of variants were not available for a variant cluster-phenotype analysis.

In our genotype-phenotype analysis comparing \textit{RPL35A} large deletions with all other pathogenic variants in \textit{RPL35A}, patients with large deletions exhibited more severe phenotypes across multiple phenotypic domains, including severity of anemia (steroid-resistance and/or RBC transfusion-dependence), neutropenia, craniofacial abnormalities, chronic GI problems, and intellectual disabilities. This suggests that a complex phenotype may be associated with the entire \textit{RPL35A} gene deletion itself or due to contiguous genes deleted in the 3q29 region. Although not well described in DBA cases caused by \textit{RPL35A} variants, similar phenotypes in DBA patients with large deletions in other ribosomal genes have been reported, including multiple congenital malformations, steroid resistant anemia, and neurodevelopmental abnormalities (e.g., \textit{RPL14}, \textit{RPS17}, \textit{RPS19}, \textit{RPS26}). Studies comparing cases with \textit{RPL35A} large deletion with DBA due to large deletions involving other ribosomal protein genes may elucidate genotype-phenotype relationships and influences of contiguous gene deletions on DBA phenotypes.

Notably, immunodeficiency and recurrent infections were both common features in cases with \textit{RPL35A} large deletions compared with all other pathogenic variants in \textit{RPL35A} ($P=0.01$ and $0.03$, respectively). Although the number of cases is small, neutropenia appears to be contributing to immunodeficiency and recurrent infections, since all these cases were also severely neutropenic and received G-CSF treatment. It is unclear from our study whether the severe neutropenia and immunodeficiency seen here is driven by \textit{RPL35A} gene deletion with or without the contribution of contiguous genes deleted in the region or by other LOF variants within \textit{RPL35A} itself, since severe neutropenia (accompanied by immunodeficiency in one case) was also seen in two of the six cases with other LOF variants.

Some of the clinical phenotypes of patients with DBA due to large deletions involving \textit{RPL35A} are similar to those observed in 3q29 deletion syndrome, including congenital abnormalities and intellectual disability, and may be influenced by additional genes deleted in the 3q29 region. Previous speculations suggested genes of importance in 3q29 deletion syndrome studies include \textit{PAK2}, \textit{DLG1}, and \textit{IL1RAP} in intellectual disabilities and \textit{TP63} and \textit{OSTN} in bone defects. However, immunodeficiency and steroid-resistant anemia are not features found in 3q29 deletion syndrome, making it less likely that a gene deleted in the 3q29 deletion syndrome region is the etiology of the immune and severe hematologic phenotype seen in our cohort. A case report of DBA caused by a \textit{RPL35A} large deletion and immune deficiency speculated that \textit{RNF168}, an autosomal recessive gene located in the 3q29 deletion syndrome region that underlies the immune deficiency Riddle Syndrome, could be a possible candidate.
leading to immunodeficiency in that case.\textsuperscript{15} However, only 12 cases in our cohort with available clinical information and a large deletion also had deletion of \textit{RNF168}. Furthermore, cases of large deletions in our study that did not overlap with \(3q29\) co-ordinates (n=4) also had evidence of steroid-resistance and/or RBC transfusion-dependence (4 of 4 cases) and neutropenia requiring treatment (3 of 4 cases).

It is possible that genes in close proximity to \textit{RPL35A} but not in the \(3q29\) deletion syndrome region may contribute to the DBA phenotypes described here. Possible candidate genes could include the closest neighbors and most frequently deleted genes \textit{LMFN} and \textit{IQCG} along with \textit{RPL35A}. These genes were deleted in 15 and 17 of 17 large deletion cases with available co-ordinates, respectively. \textit{LMFN} is a zinc-metalloprotease, and its protein product is invalidolysin.\textsuperscript{5,46} Antibodies to invalidolysin have been shown to concentrate in the edge of macrophage migration, and Drosophila mutants lacking this gene show abnormal cell migration during development.\textsuperscript{5,46} \textit{IQCG} has been observed in a somatic fusion with \textit{NUP98} in T-cell acute lymphoblastic leukemia and acute myeloid leukemia. This fusion protein has been hypothesized to block differentiation of hematopoietic stem cells.\textsuperscript{5,65} \textit{IQCG} knockout zebrafish had severely impaired levels of neutrophils, monocytes, macrophages, and lymphocytes.\textsuperscript{5,65} Further studies with functional data are warranted to determine the effect of other gene deletions on the phenotype characteristics of patients with large deletions which include \textit{RPL35A}.

In summary, we compiled the largest cohort of DBA cases with \textit{RPL35A} pathogenic variants to date. Our study is limited by small sample size, retrospective, and non-uniformity of assessment, phenotype coded as absent if not specifically mentioned in patient records, and limited family data. Since our cohort is made up of European and North American cases, the genotype-phenotype relationships may have limited generalizability. We could not examine the impact of a familial component on variation in phenotypes since we only had two families with >1 case per family. We were also limited to the 45 cases from our multi-institutional collaboration for genotype-phenotype analysis, excluding all variants found in the literature due to insufficient phenotypic data. Future studies evaluating genotype-phenotype relationships in rare diseases would benefit from detailed clinical information (as in Online Supplementary Table S4) in publications. Moreover, 30 cases included here have never been published, underscoring the importance of reporting all available cases of rare diseases to better characterize the genotype, phenotype, and potential etiology of these diseases. It is also important to include deletion analysis in molecular characterization of patients with DBA, since large deletions may not be identified on standard IBMFS gene panel testing or exome sequencing analysis.

Patients with DBA due to large deletions in \textit{RPL35A} have a complex, multi-system disease phenotype with a high frequency of hematologic as well as non-hematologic problems that is clinically different from DBA associated with other pathogenic variants in \textit{RPL35A}. Distinction of this subtype of DBA with \textit{RPL35A} haploinsufficiency due to large deletion is important for patient management, and evaluations should include thorough investigation for immunodeficiency, GI problems, developmental delays and intellectual disabilities including neuropsychiatric problems. Identification of patients with large deletions in \textit{RPL35A} should trigger early comprehensive assessment by pediatric specialists to provide optimal multidisciplinary care.

**Disclosures**

DP and JV are supported by the Czech Ministry of Health Grant 16-32105A. AK is consultant of CELGENE and Novartis and Vifor Pharma and receives Honoraria from CELGENE, Novartis and Apy Pharma. CN is a consultant and is on the advisory committee for Celgene; all other authors have no conflicts of interest.

**Contributions**

DMG and NG wrote the manuscript and performed data analyses; NG, BPA and SAS supervised the study; all authors provided data on the patients in their studies and are responsible for the reported research and have participated in the concept and design, analysis and interpretation of data, and revising the manuscript, and approved the submitted manuscript.

**Acknowledgments**

We are indebted to the participating families, whose generosity and co-operation have made this study possible.

**Funding**

This research was supported by the Intramural Research Program of the Division of Cancer Epidemiology and Genetics of the National Cancer Institute; the Czech DBA registry is supported by the Czech Ministry of Health Grant 16-32105A; we acknowledge the contributions made by Ann Carr AIS, CGC, Lisa Leathwood, RN, and Maureen Risch, RN provided through contract HHSN26120100018C with Westat Inc. (Rockville, MD, USA); this work utilized the computational resources of the NIH High Performance Computing Biowulf cluster.

**References**

9. Narla A, Vlachos A, Nathan DG. Diamond Blackfan anemia treatment: past, present,