Twelve Novel Mutations in the SLC26A3 Gene in 17 Sporadic Cases of Congenital Chloride Diarrhea

Felice Amato, Giuseppe Cardillo, Renato Liguori, Manuela Scorza, Marika Comegna, Ausilia Elce, Sonia Giordano, Laura Lucaccioni, Licia Lugi, Sabrina Cardile, Claudio Romano, Vincenza Pezzella, Giuseppe Castaldo, and Roberto Berni Canani

ABSTRACT

Objectives: We aimed to improve the knowledge of pathogenic mutations in sporadic cases of congenital chloride diarrhea (CCD) and emphasize the importance of functional studies to define the effect of novel mutations.

Methods: All member 3 of solute carrier family 26 (SLC26A3) coding regions were sequenced in 17 sporadic patients with CCD. Moreover, the minigene system was used to analyze the effect of 2 novel splicing mutations.

Results: We defined the SLC26A3 genotype of all 17 patients with CCD and identified 12 novel mutations. Using the minigene system, we confirmed the in silico prediction of a complete disruption of splicing pattern caused by 2 of these novel mutations: the c.971+3_971+4delAA and c.735+4_c.735+7delAGTA. Moreover, several prediction tools and a structure-function prediction defined the pathogenic role of 6 novel missense mutations.

Conclusions: We confirm the molecular heterogeneity of sporadic CCD adding 12 novel mutations to the list of known pathogenic mutations. Moreover, we underline the importance, for laboratories that offer molecular diagnosis and genetic counseling, to perform fast functional analysis of novel mutations.

Key Words: chloride losing diarrhea, diarrhea anion exchanger, genotype, mutations, member 3 of solute carrier family 26, splicing effect

(JPGN 2017;65: 26–30)

Congenital chloride diarrhea or chloride losing diarrhea (CCD, OMIM #214700) is a severe autosomal recessive disease characterized by watery diarrhea with a high fecal chloride loss and metabolic alkalosis. It is characterized by maternal polyhydramnios, dilated fetal bowel loops, and preterm birth. After birth, profuse watery diarrhea usually causes electrolyte imbalance with hypochloremic metabolic alkalosis or hyponatremia and life-threatening dehydration (1–3). The disease is lethal if undiagnosed or untreated, but an early diagnosis and oral NaCl/KCl replacement may improve growth and the outcome of the disease (4).

Mutation analysis is crucial for the diagnosis of such disease (5,6) that may be misdiagnosed (7–9) or confounded with a myriad of other causes of chronic diarrhea, including other genetic disorders (3,10). Furthermore, molecular diagnosis may be used for prenatal diagnosis in high-risk couples (11,12). Finally, the member 3 of solute carrier family 26 (SLC26A3) genotype influences the responsiveness to oral butyrate therapy (13–15). CCD is caused by mutations in the SLC26A3 gene (OMIM #126650), which encodes an intestinal Cl−/HCO3−, Na+-independent, exchanger (16). In ethnic groups where the disease is more frequent for founder effect, there is a single mutation (17): in Finns (incidence: 1:20–30,000), the p.V317del mutation affects up to 90% of CCD alleles; in Saudi Arabia (incidence: 1:5,000), and Kuwait (incidence: 1:3200), p.G187X affects >90% of altered chromosomes; and in Poland (incidence: 1:200,000) 50% of the CCD alleles carry the I675–7delAGTA, encoding an intestinal Cl−/HCO3−, Na+-independent, exchanger (16). In ethnic groups where the disease is more frequent for founder effect, there is a single mutation (17): in Finns (incidence: 1:20–30,000), the p.V317del mutation affects up to 90% of CCD alleles; in Saudi Arabia (incidence: 1:5,000), and Kuwait (incidence: 1:3200), p.G187X affects >90% of altered chromosomes; and in Poland (incidence: 1:200,000) 50% of the CCD alleles carry the I675–7delAGTA, encoding an intestinal Cl−/HCO3−, Na+-independent, exchanger (16).
pathogenic variant (6). Thus, data on novel mutations found in well-defined cases of CCD are contributory to improve the knowledge of pathogenic mutations. Moreover, functional studies play a key role to define the effect of novel mutations.

We analyzed by sequencing the whole-coding region of the SLC26A3 gene in 17 sporadic cases of CCD (3 of which had an affected sibling). Finally, using the in vitro functional minigene assay, we confirmed the pathogenic effect of 2 of these novel mutations.

METHODS

Subjects
The Pediatric Gastroenterology Unit at the University of Naples Federico II is an International Reference Center for patients with CCD and other forms of congenital diarrheal diseases (3,10). From 2005 to 2014, 50 cases of suspected CCD were referred to the center, and a definitive diagnosis was obtained in 38 patients. Among these, 18 were of Arabian origin and were homozygous for the G187X mutation (16) and 20 were sporadic cases of different ethnicity. Of these 20 cases of CCD, 3 included 2 affected siblings. Finally, we describe the SLC26A3 (NM_000111) genotype of 17 sporadic, unrelated cases of CCD, 7 of which are enrolled in the study on butyrate treatment (14). All the participants (guardians in the case of minors) provided written informed consent to anonymously use a DNA sample and clinical data for research purposes.

Molecular Analysis of SLC26A3 Gene

Molecular analysis was performed in the laboratory of CEINGE-Advanced Biotechnologies that acts as reference center for molecular diagnosis of inherited diseases in Campania region (approximately 6 million of inhabitants), located in southern Italy. DNA was extracted from an EDTA blood sample with the Nucleon BACC2 kit (GE Healthcare Europe GmbH, Milan, Italy). The touchdown PCR protocol that enables co-amplification of all exons under the same PCR conditions (primers and conditions are available on request). Sequencing analysis was carried out on both strands with an automated procedure (3100 Genetic Analyzer; Life Technologies). All PCR fragments were sequenced with the primers used for PCR. Novel mutations were analyzed in 100 unrelated healthy control subjects whose DNA is available c/o CEINGE-Advanced Biotechnologies that acts as reference center for molecular diagnosis of inherited diseases in Campania region (approximately 6 million of inhabitants), located in southern Italy.

RESULTS

The SLC26A3 genotype of the 17 patients with CCD is shown in (Supplemental Digital Content 1, Table, http://links.lww.com/MG/A788): 11/17 (64.7%) patients were homozygous for a SLC26A3 mutation, whereas 6/17 (35.3%) patients were compound heterozygous for 2 mutations. In the 11 homozygous patients with CCD, we tested both parents (that resulted invariably heterozygous) to exclude the possibility that a large gene deletion would cause a false homozygous mutation result; in 3 cases, in which the parents were not available for the analysis, we excluded the presence of a large deletion by quantitative RT-PCR analysis of the involved exon. Similarly, for all 6 compound heterozygous patients we studied, at least 1 parent to confirm that the 2 mutations found

www.jpgn.org
in the patients were in trans. Finally, we identified mutations in 34/34 alleles (100%). In particular, we found 21 different mutations: 9 missense, 3 nonsense, 2 single nucleotide deletion causing frame-shift, 2 splicing intronic point mutations, 1 codon duplication, 1 large deletion, 1 codon deletion, 1 intronic 4 nucleotides deletion, and 1 intronic 2 nucleotides deletion, both these latter causing a splicing effect. Among the 21 mutations, 12 (57.1%) were novel (6 missense, 3 deletions, 2 frameshift, and 1 nonsense).

All the mutations were identified by gene sequencing with the exception of the large deletion (case #1—Supplemental Digital Content 1, Table, http://links.lww.com/MPG/A788). In fact, after the gene sequencing, no gene variants were found in such patient, but exon 18 lacked to amplify. To exclude mutations in the primer(s) site, we used a more external set of primers and again, exon 18 did not amplify. Re-amplification using the long-PCR procedure revealed a homozygous large deletion; the heterozygous deletion was present in both parents. The walking primer analysis revealed the first breakpoint in position c.2008—151 and the second breakpoint in position c.2061+1546.

We functionally analyzed 2 novel intronic mutations, that is, the c.971+3_971t+4delAA and the c.735+4_c.735+7delAGTA. Both mutations were found in homozygous patients. First of all, we analyzed in silico the putative effect of both mutations, predicting a complete disruption of the physiological splicing pattern. Then, we analyzed in vitro the splicing pattern using the minigene system assay. Briefly, the exons of interest were cloned in the minigene construct (Fig. 1A) and transfected in HEI.a and T84-cells. The RT-PCR analysis on total RNA extracted from the transfected cells confirmed the in silico prediction. Indeed, in both cases, the RT-PCR analysis revealed a complete disappearance of the wild-type band (Fig. 1B and C). In particular, the sequencing analysis of each electrophoretic separated band revealed the following data: for the c.971+3_971t+4delAA mutation, we obtained 2 alternative bands: the upper one due to the skipping of exon 8 and the lower one due to the skipping of both exon 7 and exon 8. For the c.735+4_c.735+7delAGTA mutation, the mutated construct resulted in a lower band corresponding to the skipping of exon 6. We confirmed the same splicing aberration pattern in T84 cell line that expresses a splicing machinery of colon cells and represents a more physiological context for SLC26A3 gene expression.

The 6 novel missense mutations (Table 1) were not found in 200 alleles from healthy subjects. Furthermore, the novel mutations were not reported in the 1000 genomes browser (http://browser.1000genomes.org) and in the Exome Variant Server (http://evs.gs.washington.edu/EVS/). Among the missense mutations: the p.D652N involves the random coil/a-helical border changing the neutral polar serine into the apolar proline aminoacid; both the prediction tools indicated a high risk of pathogenicity. The mutation was found in 2 CCD siblings (in trans with the p.Q48X nonsense mutation). Similarly, the p.S438P involves the random coil/a-helical border changing the neutral polar serine into the apolar proline aminoacid and was found in a CCD patient that had a frameshift mutation on the other allele; both the prediction tools indicated a high risk of pathogenicity for the S438P. The A547E mutation involves the STAS domain and both the prediction tools indicated a high risk of pathogenicity. A missense mutation that abolishes the protein function in an adjacent codon (ie, the p.L496R) was already found in a CCD patient (6,18). The A547E was found in a patient compound heterozygous for the novel missense Q495P mutation. Such latter causes the change of the neutral polar glutamine with the apolar proline aminoacid. Although the PolyPhen software predicted the mutation as a possibly damaging and SIFT as a tolerated mutation, it involves the same codon of the p.L496R mutation previously found in other CCD patients (6,18,19). The p.S654P involves the STAS domain changing the neutral polar serine into the apolar proline aminoacid; the mutation was found homozygous in a CCD patient and was predicted as probably damaging by PolyPhen, whereas SIFT predicted the p.S654P as a tolerated mutation, but a mutation in an adjacent codon (ie, the p.D652N) was previously reported in patients with CCD (5,18). Finally, the p.C508R was found in a patient that had the p.R579X nonsense mutation on the other allele. The mutation changes the neutral polar cysteine with the basic polar arginine aminoacid within the random coil/intracellular domain of the protein. The mutation was predicted as probably damaging by PolyPhen, whereas SIFT predicted the p.C508R as a tolerated mutation; the p.P506Q mutation that abolishes protein function in an adjacent codon was described in CCD patients (6,17). Three other missense mutations found in our patients (ie, the p.G120S, the p.P131R, and the p.P129L), all reported as pathogenic by both the prediction tools (data not shown) had been previously found in congenital chloride diarrhea patients (6,18,19).

DISCUSSION

Molecular analysis of the SLC26A3 gene revealed the genotype of all the 17 patients with CCD and confirmed the genetic heterogeneity of sporadic CCD (6,18,19). In fact, the mutations we identified spread all over the gene, which suggests that the analysis in sporadic patients with CCD should be based on sequencing of all SLC26A3 gene coding regions, differently from ethnic groups where CCD is frequent and the analysis is routinely based on the study of a single mutation (6,17).

We identified several novel mutations clearly affecting the synthesis or the activity of the protein, and thus resulting in having a pathogenetic role. This is true for the large, 1752-bp deletion in patient #1, that removes the whole exon 18 (55 bp) likely resulting in a frameshift event (p.I670MfsX17). Similarly, the novel deletions found in patient #2 (found in 2 affected siblings) and in patient #3 are causing disease because they impair the splicing pattern, as we demonstrate in the present study. Indeed, phylogenetic analysis and in silico prediction may help to define the disease-causing effect of splicing mutations, as it is the case for both novel mutations, but only functional analysis permits to finally demonstrate such effect. In this context, the minigene assay we used represents a fast technique that allows, in approximately 2 weeks, to functionally analyze mutations with a putative splicing effect. Noteworthy, 4/21 mutations (19%) found in the present study are deletions among which one was large, confirming that such type of mutation are frequently found in monogenic disorders, as we previously demonstrated for hemophilia A, in which about 19% of 217 mutations are gene deletions (24) and in cystic fibrosis in which large rearrangements are found in approximately 3% of affected alleles (25).

Furthermore, the single nucleotide deletions observed in patients #6 and #13 (this latter found also in patient #14) cause a frameshift, and thus a truncated inactive protein. The same is true for the novel nonsense mutation found in patient #11.

More difficult it is to define the pathogenetic role of 6 novel missense mutations found in our patients with CCD. Some criteria suggest their pathogenetic role: they were absent from 200 normal alleles and in the 1000 genomes browser (http://browser.1000genomes.org) and in the Exome Variant Server (http://evs.gs.washington.edu/EVS/); in all cases, each of the parents was heterozygous for one of the mutations identified in the proband; no other mutations were identified in the SLC26A3 gene in these patients; and the amino acids affected by the mutations are highly conserved. In detail, both novel missense mutations identified in patient #15, namely, the p.Q495P and the p.A547E, lie within the mutation-bearing
hotspots characteristic of most sporadic CCD cases (19). The p.Q495P mutation could affect the correct folding of the last but 2 transmembrane domain of the protein (according to the 14 transmembrane α-helices model) (26–28), thereby altering the charge distribution. Alternatively, the mutation could alter the sulfate transporter family domain (19). The p.A547E mutation also involves the first β-sheet of the ''STAS-like'' domain. Glutamic acid is a powerful β-sheet interrupter and probably also induces a strong charge distribution change. Indeed, mutations involving the STAS-like domain have been associated to the absence of residual activity of the protein (29,30). Patient #16 is homozygous for the novel missense mutation p.S654P. The mutation involves a highly conserved amino acid belonging to the conserved loop between the third and fourth α-helix of the STAS-like domain. We suggest p.S654P is a disease-causing mutation, because the proline residue is a powerful β-sheet interrupter and mutations affecting amino acids within the “STAS-like” domain are likely to be disease causing. Moreover, the p.S654P involves a palindromic sequence (c.GACTTTTCAG) which could be more susceptible to mutations. The p.S438P mutation (patient #13)

**FIGURE 1.** A, Schematic structure of the pMGene empty and of the pMGene-DRA constructs showing the position of the novel genetic variants. B and C, RT-PCR analysis of mRNA extracted from cell transfected with the indicated pMGene constructs (M = mutated, W = wild type). The white arrows indicate the position of the PCR products with a schematic view in the right side. T84 (human lung metastase from colorectal carcinoma cells) and HeLa (cervical cancer cells).

<table>
<thead>
<tr>
<th>Mutation cDNA</th>
<th>Mutation protein</th>
<th>Domain score</th>
<th>PolyPhen prediction score</th>
<th>SIFT prediction</th>
<th>Adjacent mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>c.1181G&gt;T</td>
<td>p.S394I</td>
<td>Random coil/α-helical border</td>
<td>1 (probably damaging)</td>
<td>0 (deleterious)</td>
<td></td>
</tr>
<tr>
<td>c.1312T&gt;C</td>
<td>p.S438P</td>
<td>Random coil/α-helical border</td>
<td>0.998 (probably damaging)</td>
<td>0 (deleterious)</td>
<td></td>
</tr>
<tr>
<td>c.1640C&gt;A</td>
<td>p.A547E</td>
<td>STAS domain</td>
<td>0.993 (probably damaging)</td>
<td>0 (deleterious)</td>
<td>I544N (6,7)</td>
</tr>
<tr>
<td>c.1960T&gt;C</td>
<td>p.S654P</td>
<td>STAS domain</td>
<td>0.93 (probably damaging)</td>
<td>0.07 (tolerated)</td>
<td>D652N (6,7)</td>
</tr>
<tr>
<td>c.1522T&gt;C</td>
<td>p.C508R</td>
<td>Random coil/intracellular domain</td>
<td>0.713 (possibly damaging)</td>
<td>0.12 (tolerated)</td>
<td>P506Q (7,24)</td>
</tr>
<tr>
<td>c.1484A&gt;C</td>
<td>p.Q495P</td>
<td>Random coil/intracellular domain</td>
<td>0.696 (possibly damaging)</td>
<td>0.09 (tolerated)</td>
<td>L496R (6,7,19)</td>
</tr>
</tbody>
</table>

For each of the 6 mutations, the table reports the protein domain involved by the mutation; the PopyPhen prediction score of pathogenicity; the SIFT prediction score of pathogenicity; the list of mutations previously described involving adjacent aminoacids.
involves serine 438, the last amino acid of the random coil motif between the 11th and 12th transmembrane helix (according to the 14 transmembrane α-helices model) and proline is a powerful interrupter of α-helices. Alternatively, the mutation could alter the sulfate transporter family domain. Also in this case, the mutation lies in the hotspot region (19).

CONCLUSION

Our data confirm the high genetic heterogeneity of CCD in ethnic groups in which the disease is sporadic. This highlights the importance to analyze the whole coding region by gene sequencing in patients suspected for sporadic CCD, particularly when born to consanguineous parents. In case of novel, missense mutations, other mutations must be excluded and the novel mutation must be tested in a large number of unaffected chromosomes. Phylogenetic analysis and in silico prediction may help to confirm the disease-causing effect of novel mutations but functional analysis, here performed for 2 splicing affecting mutations, is helpful to conclude on the pathogenetic effect.

Acknowledgments: Grants from Regione Campania, DGRC 1901/09 and AIFA MRAR08W002 are gratefully acknowledged.

REFERENCES