OXFORD UNIVERSITY PRESS | Brain

SLC12A2 variants cause a neurodevelopmental disorder or cochleovestibular defect

Journal:	*Brain*
Manuscript ID	BRAIN-2019-01762.R2
Manuscript Type:	Report
Date Submitted by the Author:	05-Apr-2020
Complete List of Authors:	McNeill, Alisdair; The University of Sheffield, Sheffield Institute for Translational Neuroscience Iovino, Emanuela; University of Bologna mansard, luke; University of Montpellier Vache, Christel; University of Montpellier baux, david; University of Montpellier bedoukian, emma; CHOP cox, helen; University of Birmingham Edgbaston Campus dean, john; University of Aberdeen Goudie, David; University of Dundee kumar, ajith; Great Ormond Street Hospital for Children Newbury-Ecob, Ruth; Bristol Royal Hospital for Children Fallerini, Chiara; University of Siena lopergolo, diego; University of Siena Mari, Fransceca; University of Siena Renieri, Alessandra ; University of Siena Blanchet, Catherine; CHU MONTPELLIER, Oto-Rhino-Laryngologie et Chirurgie Cervico-Faciale willems , marjolaine; University of Montpellier roux, anne; University of Montpellier PIPPUCCI, TOMMASO; U.O. Genetica Medica - Polyclinic Sant'Orsola-Malpighi, Bologna, delpire , eric; Vanderbilt University School of Medicine
Subject category:	Genetics
To search keyword list, use whole or part words followed by an *:	GENETICS, Genetics: learning disability < GENETICS, Genetics: other < GENETICS, Whole-exome sequencing < GENETICS, Neurogenesis < SYSTEMS/DEVELOPMENT/PHYSIOLOGY, Vertigo < SYSTEMS/DEVELOPMENT/PHYSIOLOGY

SCHOLARONE™
Manuscripts

Submitted as a *Report* to *Brain*

***SLC12A2* variants cause a neurodevelopmental disorder or cochleovestibular defect**

Running title: *SLC12A2* variants in developmental disorders

Alisdair McNeill[1], Emanuela Iovino[2], Luke Mansard[3], Christel Vache[3], David Baux[3], Emma Bedoukian[4], Helen Cox[5], John Dean[6], David Goudie[7], Ajith Kumar[8], Ruth Newbury-Ecob[9], Chiara Fallerini[10,11], Alessandra Renieri[10,11], Diego Lopergolo[10,11], Francesca Mari[10,11], Catherine Blanchet[12] Marjolaine Willems[13], Anne-Francoise Roux[3],Tommaso Pippucci[14], and Eric Delpire[15]

1. Department of Neuroscience, University of Sheffield, Sheffield, S10 2HQ, UK; Sheffield Clinical Genetics Service, Sheffield Children's Hospital NHS Foundation Trust.
a.mcneill@sheffield.ac.uk

2. Department of Medical and Surgical Sciences, University of Bologna, Bologna, Italy.

3. Laboratory of Molecular Genetics, CHU Montpellier, University of Montpellier, Montpellier, France

4. Roberts Individualized Medical Genetics center, Children's Hospital of Philadelphia, Philadelphia. PA 19104.

5. Regional Clinical Genetics Unit, Birmingham Women's and Children's Hospital NHS Foundation Trust, Mindelsohn Way, Birmingham. B15 2 TG.

6. North of Scotland Genetics Service, Aberdeen Royal Infirmary, Foresterhill, Aberdeen. AB25 2ZA.

7. East of Scotland Regional Genetics Service, Level 6, Ninewells Hospital, Dundee. DD2 1SY.

8. Clinical Genetics Unit, Great Ormond Street Hospital, Great Ormond Street, London WC1N 3JH

9. Bristol Regional Genetics Service, St Michael's Hospital, Southwell Street, Bristol. BS2 8EG.

10. Medical Genetics, University of Siena, Siena, Italy

11. Genetica Medica, Azienda Ospedaliera Universitaria Senese, Siena, Italy

12. Centre of Reference for Genetic Sensory diseases, CHU Montpellier, University of Montpellier, Montpellier, France

13. Department of Clinical Genetics, CHU Montpellier, University of Montpellier, Montpellier, France

14. Medical Genetics Unit, Polyclinic Sant'Orsola-Malpighi University Hospital, Bologna, Italy.

15. Department of Anesthesiology, Vanderbilt University School of Medicine, Nashville, TN,
37232, USA. eric.delpire@vanderbilt.edu

Corresponding authors:

Alisdair McNeill, PhD FRCP Edin DCH   Eric Delpire, Ph.D.

Department of Neuroscience   Department of Anesthesiology

University of Sheffield   Vanderbilt University Medical Center

385a Glossop Road   T-4202 MCN, 1161 21st Avenue South

Sheffield, S10 2HQ   Nashville, TN 37232-2520, USA

Tel: +44 (0)114 22 22267   Tel: +1-615-343-7409

a.mcneill@sheffield.ac.uk   eric.delpire@vanderbilt.edu

Abstract

The *SLC12* gene family consists of *SLC12A1 - SLC12A9*, encoding electroneutral cation-coupled chloride cotransporters. *SCL12A2* has been shown to play a role in corticogenesis and therefore represents a strong candidate neurodevelopmental disorder gene. Through trio exome sequencing we identified *de novo* mutations in *SLC12A2* in 6 children with neurodevelopmental disorders. All had developmental delay or intellectual disability ranging from mild to severe. Two had sensorineural deafness. We also identified *SLC12A2* variants in 3 individuals with non-syndromic bilateral sensorineural hearing loss and vestibular areflexia. The *SLC12A2 de novo* mutation rate was demonstrated to be significantly elevated in the Deciphering Developmental Disorders cohort. All tested variants were shown to reduce cotransporter function in *Xenopus laevis* oocytes. Analysis of *SLC12A2* expression in fetal brain at 16-18 weeks post conception revealed high expression in radial glial cells, compatible with a role in neurogenesis. Gene co-expression analysis in cells robustly expressing *SLC12A2* at 16-18 weeks post conception identified a transcriptomic program associated with active neurogenesis. We identify *SLC12A2 de novo* mutations as the cause of a novel neurodevelopmental disorder and bilateral non-syndromic sensorineural hearing loss and provide further data supporting a role for this gene in human neurodevelopment.

Keywords

Corticogenesis, neurodevelopmental disorder, *de novo* mutation, exome, brain.

Introduction

Neurodevelopmental disorders (NDD) affect 1-5% of the population and demonstrate clinical and aetiological heterogeneity (Wright *et al.*, 2015). *De novo* mutations (DNM) are associated with NDD in around 25% of cases (McRae *et al.*, 2017). Corticogenesis is the process by which neuronal progenitors proliferate and migrate to form the cerebral cortex (Urbán and Guillemot, 2014). Corticogenesis is tightly regulated by transcriptional programs, with temporal and spatial regulation of gene expression. In humans corticogenesis begins around embryonic day 42 and is largely completed by birth (Urbán and Guillemot, 2014). Many of the genes affected by DNM in NDD play a role in corticogenesis (McRae *et al.*, 2017; Zawerton *et al.*, 2019), and corticogenesis genes are strong candidates for NDD. The *SLC12* gene family consists of *SLC12A1 - SLC12A9*, encoding electroneutral cation-coupled chloride transporters (Arroyo *et al.*, 2013). Several of these genes are known to cause human disease (Table S1). There is no clear gene-disease relationship for *SLC12A2* (NKCC1). However, *SCL12A2* has been shown to play a role in corticogenesis (Young *et al.*, 2012) and the excitatory-inhibitory GABA switch during brain development (He *et al.*, 2014). *SLC12A2* represents a good candidate gene for NDD.

Study of an *Slc12a2*/NKCC1 null mouse demonstrated that the transporter is a key mechanism in the accumulation of the K^+-rich endolymph in the inner ear (Delpire *et al.*, 1999). Absence of NKCC1 causes sensorineural deafness and balance deficits. Little is known about a brain phenotype in the NKCC1 knockout mouse. However, NKCC1 expression in central neurons is developmentally regulated in rodents (Plotkin *et al.*, 1997). Expression is likely highest when the neurons are born in the subventricular zone at a time when the intracellular chloride (Cl^-) concentration is highest (Ben-Ari, 2012). High Cl^- facilitates the development of GABA-mediated giant synaptic potentials, by which GABA excites developing neurons to promote growth and synapse formation when glutamatergic inputs have not yet developed (Ben-Ari, 2012). Disruption of NKCC1 expression during development is likely to have consequences for brain development.

SLC12A2 undergoes alternative splicing, with 8-isoforms identified by the Genotype-Tissue Expression Project (GTEx)(Figure S1)(Stranger *et al.*, 2017). The full-length isoform is 27 exons, with exon 21 being spliced out in the other major isoform. In mouse cochlea only the exon 21 containing isoform is expressed (Dixon *et al.*, 1999). Deafness in the *sy* mouse is due to an exon 21 frameshift in *Slc12a2* (Dixon *et al.*, 1999). In the brain both exon 21 containing and exon 21 deleted isoforms are expressed (Morita *et al.*, 2014). Here, we describe 6 patients with DNM in

SLC12A2 associated with a NDD and 3 with bilateral sensorineural hearing loss (BLSNHL) and exon 21 variants.

Methods

Ascertainment of individuals with *SCL12A2* variants

Probands with protein altering SNVs in *SLC12A2* were identified in patients 1-5 by exome sequencing in the Deciphering Developmental Disorders study (DDD)(Wright *et al.*, 2015). Subjects 6-9 were identified via Genematcher (Sobreira *et al.*, 2015). Written consent was obtained from parents/guardians.

***In silico* assessment of pathogenicity of *SCL12A2* variants**

The effects of *SLC12A2* variants were assessed using multiple *in silico* tools (Schwarz *et al.*, 2010; Shihab *et al.*, 2013). The excess in occurrence of *de novo SLC12A2* protein-altering variants in DDD was assessed using DenovolyzeR(Ware *et al.*, 2015), which compares the observed:expected number DNMs to identify an elevated de novo mutation rate for a given gene. Spatial clustering of missense *SLC12A2* DNMs was assessed using Denovonear, which calculates the probability of the observed spatial pattern of DNM arising by chance. We used MuPiT, which maps the effect of SNVs onto 3D-protein structures, to visualise the structural effects of DNMs (Niknafs *et al.*, 2013).

Transcriptomic study of *SLC12A2* expression in developing human brain

Variations in *SLC12A2* transcript levels in the human brain among different developmental stages and anatomical regions were investigated using data from the BrainSpan Atlas of the Developing Human Brain (Miller *et al.*, 2014). Single cell RNA (scRNA) sequencing data from the human brain at 16-18 weeks gestation was obtained from (Pollen *et al.*, 2015). NetworkAnalyst, which uses protein-protein interaction networks to analyse gene expression studies, was used to explore transcriptomic differences in scRNA data between *SLC12A2* expressing and non-expressing cells (Xia *et al.*, 2015).

***In silico* study of expression of *SLC12A2* splice isoforms**

RNA-seq data from the BrainSpan Atlas of the Developing Human Brain was used to analyse levels of *SLC12A2* exon expression during brain development. RNA-seq data from developing mouse cochlea was examined for alternative splicing of *SLC12A2* during cochlear development (Ranum *et al.*, 2019).

***In vitro* assessment of NKCC1 ion transporter function**

NKCC1 function was assessed through K^+ influx measurements into *Xenopus laevis* oocytes injected with wild-type or mutant NKCC1 cRNA. Detailed protocols have been published and summarised in the supplementary methods (Delpire *et al.*, 2011).

Results

Identification of *SLC12A2* variants in children with neurodevelopmental disorders

Through trio exome sequencing, we identified 6 children with NDD and DNM in *SLC12A2*. Detailed clinical reports are in the supplementary material and Table 1. All had intellectual disability or developmental delay varying from mild to severe. Three had an autistic spectrum disorder. Patient 1 had cerebral cortical dysplasia on brain MRI. Patient 5 agenesis of the corpus callosum. Two had BLSNHL. Patient 4 had spastic diplegia and patient 5 spastic quadriparesis. In addition we identified 3 individuals from 2 unrelated families with congenital non-syndromic BLSNHL and an *SCL12A2* variant (Table 1). There was no shared facial dysmorphology. None had a pathogenic copy number variant, any additional candidate DNM (Table S2) or causal variant in a deafness gene.

In silico assessment of *SLC12A2* variants supports pathogenicity

Seven of the 8 identified *SLC12A2* variants were *de novo* and all had CADD scores > 20 (Table 1, Figure 1a). The missense variants affected evolutionarily conserved residues, conserved between *SLC12* gene family members (Figure S2). We assessed whether the DNM rate in the DDD cohort was elevated. We collected all 8 *SLC12A2* protein-altering DNMs (the 5 pathogenic variants plus 3 variants predicted to be benign, table in data supplement) in DDD. Using DenovolyzeR, we showed a 9.62-fold DNM enrichment in *SLC12A2* (p= 2.71e-06, after correction for multiple testing p= 0.05).

SLC12A2 sequence data from the gnomAD database demonstrated significant constraint for PTV in *SLC12A2* (58.7 expected versus 11 observed, pLI = 0.96, observed:expected 0.19 [0.12-0.31]), but not for missense SNVs (594 observed versus 430 expected, Z = 2.4, observed:expected 0.72 [0.67-0.78]). We then used gnomAD data to examine the spatial distribution of missense SNVs in *SLC12A2*, to look for differential prevalence of missense variants in protein domains (implying regional constraint). There was no difference in the percentage of residues with a synonymous variant in the functionally important domains (transmembrane domains)(Chew *et al*, 2019) and domains of no known function (19% vs 22%, p = 0.13). In the functionally significant domains, 35% of residues had a missense variant compared to 43% of residues in the domains of no known functional significance (chi-squared, p = 0.0037). In addition, constraint data from (Samocha *et al*., 2017) indicates that amino acids 1 – 836 are depleted of missense variants (observed 173, expected

294.8, ratio 0.58, chi squared 50.3) while the remainder of the protein is not. This suggests constraint of missense variation in the functionally significant domains of *SLC12A2*.

Pathogenic *SLC12A2* variants alter NKCC1 activity

To assess the functional significance of the mutations, we injected wild-type NKCC1 and mutant NKCC1 cRNAs in *Xenopus laevis* oocytes and performed standard K$^+$ influx measurements under isotonic (basal) and hypertonic (stimulated) conditions. We used the mouse cDNA, which is 95% conserved compared to human *SLC12A2*. For every mutation tested, Figure S2 demonstrates homology of the mutated and surrounding residues within the *Slc12a* transporter family and among NKCC1 proteins from 6 species (sea urchin and five vertebrates). All tested mutations demonstrated significant reduction in K$^+$ influx (Figure 2, detailed results in supplementary material).

Spatial proximity of *SLC12A2* missense DNMs altering NKCC1 activity

We wondered whether DNMs (p.(Ala327Val), p.(Arg410Gln), p.(Asn376Ile), p.(Ala379Pro), p.(Glu980Lys)) clustered within the protein. Clustering analysis with Denovonear demonstrated a greater spatial proximity within the protein than expected by chance (p= 0.027), with the majority of DNMs falling into transmembrane domains (Figure 1b). We then used MuPiT to map these DNMs onto NKCC1 three-dimensional protein structure, showing all DNMs embedded within the transmembrane core of the protein (Figure 1c).

Expression of *SLC12A2* in developing human brain supports a role in neurogenesis

We next examined *SLC12A2* expression in developing human brain using microarray data from the BrainSpan atlas. *SLC12A2* expression was significantly higher (Mann-Whitney U-test, Z= -8.36 p < 0.001) in neuroanatomical regions with high neurogenesis (ventricular zone, subventricular zone) compared with regions of less active neurogenesis (cortical plate, subplate) at 15-16 weeks gestation (Figure 3a). For a range of brain expressed genes with no known role in neurogenesis there was no differential expression in neurogenic niches (data not shown).

Single-cell RNA sequencing data from 16-18 weeks gestation fetal brain was then used to identify populations of cells which express *SLC12A2* during development. The marker genes defined in Pollen *et al* were used to identify radial glia (e.g. *PAX6, SOX2, VIM*), intermediate progenitors (e.g. *EOMES, RBFOX1*) and neurons (e.g. *MYT1L, NEUROD1, SATB2*). *SLC12A2* expression was significantly higher in radial glia than in intermediate progenitors (mann-whitney u-test, Z = -3.3, p

= 0.001) or neurons (z = -5.2, p < 0.001), but did not differ between intermediate progenitors and neurons (z = -1.06, p = 0.28, Figure 3b). In radial glial there was no expression of other *SLC12* gene family members known to have a role in neuronal development or function. This suggests other *SLC12A2* family members cannot compensate for reduction in *SLC12A2* function associated with *SLC12A2* DNM.

In silico **study of expression of *SLC12A2* splice isoforms**

RNA-seq data from BrainSpan quantifies levels of exon expression. Figure S3a demonstrates an excess of transcripts for exon 1-28 compared to exon 21 in human fetal brain. *SLC12A2* isoforms lacking exon 21 therefore exist in developing human brain. RNA-seq of developing mouse cochlea demonstrated no alternative splicing of exon 21; the exon 21 containing isoform is the only transcript expressed (Figure S3b).

SLC12A2 **expressing cells display transcriptomic profiles of active neurogenesis**

To investigate the functional properties of *SLC12A2* expressing cells we used the single cell RNA sequencing data from Pollen *et al* to select cells robustly expressing *SLC12A2* (>100 counts per million) and those with no *SLC12A2* expression. This was done agnostic to cell type. We then used NetworkAnalyst to identify differentially expressed genes in *SLC12A2* expressing cells. This identified 589 differentially expressed genes (corrected p-value < 0.05 with > 2 fold difference in expression). Images of networks generated and a full list of enrichment terms is in the supplemental data (Figure S4).

Given that we were analysing scRNA data from fetal brain, we used tissue specific (brain cortex) co-expression network analysis to form a network from differentially expressed genes. NetworkAnalyst created a gene co-expression network (supplementary data). This network was enriched for Reactome terms (e.g. L1CAM interaction and axon guidance [both FDR corrected p = 0.00085]) and Gene Ontology (GO) biological process terms (e.g. axonogenesis [p = 0.0027], neurodevelopment [p = 0.0036]) relevant to neurodevelopment. We then used the label propagation algorithm to identify 5 modules within the network (supplementary data). Details of the enrichment analysis of all 5 modules is in the data supplement. Module 1 (coloured blue) was the largest (435 genes, Figure S4a). It was enriched for Reactome term L1CAM interaction (p = 0.010) and GO biological process terms cytoskeleton dependent intracellular transport (p = 0.00005) and synaptic transmission (p = 0.00072) as well as Panther biological process term nervous system development (p = 0.017). Module 2 (coloured red, 104 genes, Figure 4b) was enriched for the Reactome term

notch-HLH transcription pathway (p = 0.0016) and GO biological process terms for neuron formation (neuron differentiation, generation of neurons, neurogenesis, nervous system development, all p=0.01). <u>Modules 3 (coloured white, 123 genes), 4 (coloured green, 117 genes) and 5 (coloured yellow, 123 genes) were not enriched for neurogenesis terms</u>. *SCL12A2* expressing cells manifest a transcriptomic program suggesting an active role in neurogenesis.

Discussion

We describe 6 individuals with a pleiotropic NDD associated with DNM in *SCL12A2*. All had developmental delay, ranging from mild to severe. 3 had autistic spectrum disorder. In addition we identified 3 patients with BLSNHL and no NDD. Two individuals with likely pathogenic variants in *SLC12A2* have previously been reported. We described a female patient with a PTV in exon 22 of *SLC12A2* (Delpire *et al.*, 2016). She presented with lung, gastrointestinal tract, endocrine and exocrine gland deficits with seizure like episodes and EEG abnormalities, but no hearing impairment. Recently, a 5-year old boy with uniparental disomy for chromosome 5 with a 22 kb deletion of *SLC12A2* was described with bilateral sensorineural deafness, global developmental delay and failure to thrive (Macnamara *et al.*, 2019). Our report confirms that *SLC12A2* variants are associated with a pleiotropic NDD and exon 21 variants with BLSNHL.

Pleiotropy is well recognised in NDD (McRae *et al.*, 2017). The explanation for pleiotropy associated with *SLC12A2* is unclear. The severe phenotype in the boy reported by MacNamara *et al* is likely related to the presence of a homozygous variant in *SLC12A2,* combined with the effects of uniparental disomy for chromosome 5. Patients 3 and 5 in our series had PTV and their phenotypes were subjectively more severe than those of the patients with missense variants. A larger series of *SLC12A2* variants will be required to confirm any genotype-phenotype correlation. Patient 7 was 47,XYY. The contribution of this is not clear. XYY-syndrome has been associated with a mild reduction in intelligence quotient (IQ)(Green *et al.*, 2019). There is no clear association between BLSNHL or tongue fasciculations and XYY-syndrome. The cochlea expresses only the *SLC12A2* isoform containing exon 21, suggesting exon 21 plays a critical developmental role in cochlea (Dixon *et al.*, 1999) and explaining why exon 21 variants cause BLSNHL. Notably, deafness in the *sy* mouse is due to an exon 21 *Slc12a2* mutation (Dixon *et al.*, 1999). In human fetal brain, significant amounts of the *SLC12A2* isoform lacking exon 21 are expressed (Morita *et al.*, 2014). This may compensate for deleterious effects of exon 21 variants and explain why exon 21 variants do not cause a NDD.

In both murine models (Magalhães and Rivera, 2016) and cell systems (Young *et al.*, 2012), loss of *Slc12a2*/NKCC1 has been shown to inhibit neurogenesis. *SLC12A2* is highly expressed in areas of active neurogenesis and *SLC12A2* expression is higher in radial glia than in intermediate progenitor cells or neurons at 16-18 weeks gestation. Transcriptomic profiling indicates that *SLC12A2* expressing cells (at 16-18 weeks gestation) manifest a transcriptomic program reflecting active neurogenesis. Reduced function of *SLC12A2* might result in a NDD by altering the delicate process of corticogenesis, and/or by dysregulating the excitatory-inhibitory GABA switch.

The phenotype of humans with *SLC12A2* variants has similarities to animal models. An slc12a2 null zebrafish displays collapse of the otic vesicle with reduced endolymph (Abbas and Whitfield, 2009). This paper did not describe the brain phenotype. The otic malformation in the zebrafish model has relevance to the hearing loss in *SLC12A2* variant carriers. Several *Slc12a2* murine models exist. We initially reported an *Slc12A2* null mouse with cochlear malformations, loss of hair cells and hearing impairment (Delpire *et al.*, 1999). The brain phenotype in this mouse has not been studied extensively. Reduced neuronal proliferation has been demonstrated in the lateral ganglionic eminence of the null mouse (Magalhães and Rivera, 2016), and NKCC1 knockdown by short hairpin RNA reduced neuronal proliferation in the murine subventricular zone (Young *et al.*, 2012). The similarities between animal models and the human phenotype of *SCL12A2*, the role of *SLC12A2* in neuronal development, the results of *in silico* analyses and the effects of the SNVs on NKCC1 biochemical function provide strong support for a causal role of *SLC12A2* variants in NDD and BSNHL.

Acknowledgements

Competing interests

The authors declare there are no competing interests.

Internet resources

References

Abbas L, Whitfield TT. Nkcc1 (Slc12a2) is required for the regulation of endolymph volume in the otic vesicle and swim bladder volume in the zebrafish larva. Development 2009; 136: 2837–2848.

Arroyo JP, Kahle KT, Gamba G. The SLC12 family of electroneutral cation-coupled chloride cotransporters. Mol Aspects Med 2013; 34: 288–298.

Ben-Ari Y. The yin and yen of GABA in brain development and operation in health and disease. Front Cell Neurosci 2012; 6: 45

Chen TA, Orlando BJ, Zhang J, Latorraca NR, Wang A, Hollingsworth SA, et al. Structure and mechanism of the cation-chloride cotransporter NKCC1. Nature 2019; 572: 488-492.

Delpire E, Gagnon KB, Ledford JJ, Wallace JM. Housing and husbandry of Xenopus laevis affect the quality of oocytes for heterologous expression studies. J Am Assoc Lab Anim Sci 2011; 50: 46–53.

Delpire E, Lu J, England R, Dull C, Thorne T. Deafness and imbalance associated with inactivation of the secretory Na- K-2Cl co-transporter. Nat Genet 1999; 22: 192–195.

Delpire E, Wolfe L, Flores B, Koumangoye R, Schornak CC, Omer S, et al. A patient with multisystem dysfunction carries a truncation mutation in human SLC12A2, the gene encoding the

Na-K-2Cl cotransporter, NKCC1. Cold Spring Harb Mol case Stud 2016; 2: a001289.

Dixon MJ, Gazzard J, Chaudhry SS, Sampson N, Schulte BA, Steel KP. Mutation of the Na-K-Cl co-transporter gene Slc12a2 results in deafness in mice. Hum Mol Genet 1999; 8: 1579–1584.

Green T, Flash S, Reiss AL. Sex differences in psychiatric disorders: what we can learn from sex chromosome aneuploidies. Neuropsychopharmacology 2019; 44: 9–21.

He Q, Nomura T, Xu J, Contractor A. The developmental switch in GABA polarity is delayed in fragile X mice. J Neurosci 2014; 34: 446–450.

Macnamara EF, Koehler AE, D'Souza P, Estwick T, Lee P, Vezina G, et al. Kilquist syndrome: A novel syndromic hearing loss disorder caused by homozygous deletion of SLC12A2. Hum Mutat 2019; 40: 532–538.

Magalhães AC, Rivera C. NKCC1-Deficiency Results in Abnormal Proliferation of Neural Progenitor Cells of the Lateral Ganglionic Eminence. Front Cell Neurosci 2016; 10: 200.

McRae JF, Clayton S, Fitzgerald TW, Kaplanis J, Prigmore E, Rajan D, et al. Prevalence and architecture of de novo mutations in developmental disorders. Nature 2017; 542: 433–438.

Miller JA, Ding S-L, Sunkin SM, Smith KA, Ng L, Szafer A, et al. Transcriptional landscape of the prenatal human brain. Nature 2014; 508: 199–206.

Morita Y, Callicott JH, Testa LR, Mighdoll MI, Dickinson D, Chen Q, et al. Characteristics of the cation cotransporter NKCC1 in human brain: Alternate transcripts, expression in development, and potential relationships to brain function and Schizophrenia. J Neurosci 2014; 34: 4929–4940.

Niknafs N, Kim D, Kim R, Diekhans M, Ryan M, Stenson PD, et al. MuPIT interactive: webserver for mapping variant positions to annotated, interactive 3D structures. Hum Genet 2013; 132: 1235–43.

Plotkin MD, Snyder EY, Hebert SC, Delpire E. Expression of the Na-K-2Cl cotransporter is developmentally regulated in postnatal rat brains: a possible mechanism underlying GABA's excitatory role in immature brain. J Neurobiol 1997; 33: 781–95.

Pollen AA, Nowakowski TJ, Chen J, Retallack H, Sandoval-Espinosa C, Nicholas CR, et al. Molecular Identity of Human Outer Radial Glia during Cortical Development. Cell 2015; 163: 55–67.

Ranum PT, Goodwin AT, Yoshimura H, Kolbe DL, Walls WD, Koh JY, et al. Insights into the Biology of Hearing and Deafness Revealed by Single-Cell RNA Sequencing. Cell Rep 2019; 26: 3160-3171.e3.

Samocha KE, Kosmicki JA, Karczewski KJ, O'Donnell-Luria AH, Pierce-Hoffman E, MacArthur DG, et al. Regional missense constraint improves variant deleteriousness prediction. Biorxiv. doi.org/10.1101/148353

Schwarz JM, Rödelsperger C, Schuelke M, Seelow D. MutationTaster evaluates disease-causing potential of sequence alterations. Nat Methods 2010; 7: 575–6.

Shihab HA, Gough J, Cooper DN, Stenson PD, Barker GLA, Edwards KJ, et al. Predicting the functional, molecular, and phenotypic consequences of amino acid substitutions using hidden Markov models. Hum Mutat 2013; 34: 57–65.

Sobreira N, Schiettecatte F, Valle D, Hamosh A. GeneMatcher: a matching tool for connecting investigators with an interest in the same gene. Hum Mutat 2015; 36: 928–30.

Stranger BE, Brigham LE, Hasz R, Hunter M, Johns C, Johnson M, et al. Enhancing GTEx by bridging the gaps between genotype, gene expression, and disease. Nat Genet 2017; 49: 1664–1670.

Urbán N, Guillemot F. Neurogenesis in the embryonic and adult brain: Same regulators, different roles. Front Cell Neurosci 2014; 8: 396

Ware JS, Samocha KE, Homsy J, Daly MJ. Interpreting de novo Variation in Human Disease Using denovolyzeR. Curr Protoc Hum Genet 2015; 87: 7.25.1-7.25.15.

Wright CF, Fitzgerald TW, Jones WD, Clayton S, McRae JF, Van Kogelenberg M, et al. Genetic diagnosis of developmental disorders in the DDD study: A scalable analysis of genome-wide research data. Lancet 2015; 385: 1305–1314.

Xia J, Gill EE, Hancock REW. NetworkAnalyst for statistical, visual and network-based meta-analysis of gene expression data. Nat Protoc 2015; 10: 823–844.

Young SZ, Morgan Taylor M, Wu S, Ikeda-Matsuo Y, Kubera C, Bordey A. NKCC1 knockdown decreases neuron production through GABAA-regulated neural progenitor proliferation and delays dendrite development. J Neurosci 2012; 32: 13630–13638.

Zawerton A, Yao B, Yeager JP, Pippucci T, Haseeb A, Smith JD, et al. De Novo SOX4 Variants Cause a Neurodevelopmental Disease Associated with Mild Dysmorphism. Am J Hum Genet 2019; 104: 246–259.

Figure legends

Figure 1. *SLC12A2* missense variants.

A. Chart summarising frequency of clinical features in children with *SLC12A2 de novo* mutations and a NDD. **B**. Schematic representation of NKCC1 with 12 transmembrane domains and inverted 5 TMs + 5 TMs symmetry (shaded TMs), followed by TM11 and TM12 (white TMs). The position of patient mutations is indicated in red. **C**. Three dimensional structure of NKCC1 demonstrating location of missense variants at core of protein, with high potential to disrupt protein structure.

Figure 2. **Functional analysis of NKCC1 mutation in *Xenopus laevis* oocytes**. **A-B**. Box plots demonstrating K^+ influx measured in oocytes injected with water (negative control), water containing 15 ng wild-type (control) or mutant mouse NKCC1 cRNA. Horizontal line represents median, extent of box demonstrates interquartile range and whiskers minimum and maximum. K^+ fluxes were measured under basal isosmotic (200 mOsM) or stimulated hyperosmotic (270 mOsM) conditions and are expressed in pmoles K^+ per oocyte per hour. As the mutants were tested in different experiments on different oocytes, each mutant flux is accompanied by its own control flux. Statistical differences were determined using one way ANOVA. Each experiment represents 20-25 oocytes per mutation. Note that the mouse residue numbers are slightly different from human.

Figure 3. Transcriptomic analysis of *SCL12A2* expression in developing human brain.

Figure 3. Transcriptomic analysis of SCL12A2 expression in developing human brain. 3a. Box plots of microarray data comparing SLC12A2 expression in neuroanatomical regions of high and less active neurogenesis at 15 weeks gestation. Horizontal lines represent median, extent of box interquartile range and whiskers minimum and maximum. Expression is significantly higher in areas of active neurogenesis (*). 3b. Single cell RNA sequencing data demonstrating significantly higher expression of SLC12A2 in radial glia cells (RGC) compared to intermediate progenitor cells (IPC)(mann-whitney u-test, Z = -3.3, p = 0.001) and neurons (z = -5.2, p < 0.001). Units are counts per million reads (CPM).

Figure S1. *SLC12A2* **splice isoforms.**

Eight *SLC12A2* isforms are identified by GTEx https://www.gtexportal.org/home/gene/SLC12A2. Note canonical isoform ENST00000262461.6 has 27 exons. ENST00000343225.4 has exon 21 spliced out and represents the minor isoform found in the developing brain.

Figure S2. **Sequence alignment of** *SLC12* **gene family members.**

The amino acid sequences of NKCC1 from eight species were aligned using VectorNti Suite 6.0 (Invitrogen). Portions representing the regions of interest were selected as follows: in **A**, a portion of TMD2 highlighting the conservation of the Ala327 residue; in **B**, a portion of TMD4 showing conservation of Arg410; in **C**, amino acids within a portion of the carboxyl-terminus showing conservation of Trp892; in **D**, portion of TM3 showing conservation of residues Asn376 and Ala377; and in E, a portion of the carboxyl-terminus showing some degree of conservation of residues Glu979 and Glu980. To assess conservation between the different members of the SLC12A gene products, the amino acids from human NKCC1, NKCC2, NCC, and KCC1-4 were also aligned. In **F**, residue Ala327 is less conserved among the different cotransporters; in **G**, Arg410 is highly conserved; in **H**, Trp758 is also a highly conserved residue; and in **I**, Asn376 and Ala377 are only conserved within the Na$^+$-dependent cotransporters. Identical residues across are highlighted by yellow background, whereas conserved substitutions are shown in green or blue background. Non-conserved residues are shown in white background.

Figure S3. Expression of alternatively spliced exons and isoforms of *SCL12A2*.

a. The ratio of exon 21 expression to expression of all other 27 exons of *SLC12A2* is <1.0 in the first and second trimester of pregnancy. This demonstrates the presence of transcripts which do not contain exon 21.

b. Image captured from https://morlscrnaseq.org/ an atlas of long-read single cell RNA expression data from murine inner ear. Sashimi plot demonstrates no alternative splicing events at exon 21. The only transcript expressed in mouse cochlea therefore contains exon 21.

Figure S4. Gene co-expression network analysis of single cell RNA data from *SLC12A2* expression cells.

    A. Figure shows module 1 identified by gene co-expression network based on differentially expressed genes created by NetWorkAnalyst.   The degree filter was used so that only hub genes are shown in this figure.

    B. Table of significantly enriched terms for whole network.

Table 1. Summary of clinical and genomic data for *SLC12A2* variant carriers.

	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6
Age, sex	1-year, M	9- years, M	15-years, F	3-years, F	6-years, F	21 years, M
Variant (Genome build GRCh-37/hg19)	g. 5 127450305C>T	g. 5 127469897G>A	g. 5 127503511G>A	g. 5 127466837A>T	g. 5 127420201dup	g. 5 127466845delinsCT
c.DNA variant NM_001046.3	c.980C>T	c.1229G>A	c.2675G>A	c.1127A>T	c.555dupG	c.1135_1136delGCinsCT
ACMG criteria	4	4	5	4	4	4
Inheritance	*De novo*	*De novo*	*De novo*	*De novo*	*De novo*	*De novo*
Amino acid change	p.(Ala327Val)	p.(Arg410Gln)	p.(Trp892*)	p.(Asn376Ile)	p.(His186AlafsTer17)	p.(Ala379Leu)
Exon (of 27)	4	6	18	5	1	5
Gnomad frequency	0	0.0000070 (2 cases)	0	0	0	0
CADD-Phred score	26.0	31	40	27.3	33	28
GERP	5.36	4.94	4.8	4.92	3.97	4.92
MutationTaster	Disease causing	Disease causing	Disease causing (1)	Disease causing	N/A	Disease causing (1)

	(1)	(1)		(1)		
FATHMM	Damaging (-5.14)	Damaging (-2.17)	Damaging (0.977) (FATHMM-MKL)	Damaging (-5.18)	N/A	Damaging (-5.03)
PolyPhen-2	Probably Damaging (1)	Probably Damaging (1)	N/A	Probably Damaging (1)	N/A	Probably Damaging (1)
Growth	OFC 40 cm (1st) Weight 3kg (64th)	OFC 52cm (6th) weight 29.8kg (39th) Height 139 cm (59th)	OFC 55cm (46th) Height 165 cm (65th)	OFC 51.8 cm (82nd)	Weight 26 kg (2nd - 9th) OFC 46 cm (<0.4th)	OFC 59 cm Height 174 cm (36th) Weight 64kg (27th)
Neurology						
Brain MRI	Cortical dysplasia	N/A	N/A	Normal	Agenesis of corpus callosum	N/A
Development	Not sitting Non-verbal		Walking 3 years non-verbal	Delayed walking and speech	Non-ambulant Non-verbal	Walked 13 months Speech 6 years
Intellectual disability	Severe	Mild	Severe	None	Severe	Severe
Autism	No	Yes	Yes	No	No	Yes
ADHD	No	No	No	No	No	No
Epilepsy	No	No	No	No	No	No
Motor	N/A	N/A	Stereotypies Hypotonia	Spastic diplegia	Spastic quadriplegia	

Hearing	BLSNHL	Not assessed	BLSNHL	Normal	Normal	Normal
Ocular	Unilateral iris coloboma	No	No	No	No	No
Cardiac Defect	Ventricular septal defect	No	No	No	No	No
Gastrointestinal	Tracheo-esophageal fistula	No	Reflux	No	Unsafe swallow, gastrostomy feed	No

	Case 7	Case 8	Case 9	Case 10 (Macnamara)	Case 11 (Delpire)
Age, sex	2-years, M	44-years, M	5-years, M	5-years, M	17-years, F
Variant (Genome build GRCh-37/hg19)	g. 5 127512805G>A	g. 5 127512802G>A	g. 5 127512802G>A	Uniparental disomy 5, 22kb deletion *SLC12A2*	g. 5 127514355_127514365del
cDNA NM_001046.3	c.2938G>A	c.2935G>A	c.2935G>A	N/A	c.3076_3086del
ACMG	4	4	4	N/A	5
Inheritance	*De novo*	Unknown	Inherited from case 8	Uniparental disomy	*De novo*
Amino acid change	p.(Glu980Lys)	p.(Glu979Lys)	p.(Glu979Lys)	N/A	p.(Val1026PhefsTer2)
Exon (of 28)	21	21	21	2-7	22
Gnomad frequency	0	0	0	0	0
CADD-Phred score	23	23	23	N/A	36
GERP	5.11	5.11	5.11	N/A	4.96
MutationTaster	Disease causing (1)	Disease causing (1)	Disease causing (1)	N/A	N/A

FATHMM	Damaging (-1.91)	Damaging (-1.9)	Damaging (-1.9)	N/A	N/A
PolyPhen-2	Possibly damaging (0.682)	Possibly damaging (0.799)	Possibly damaging (0.799)	N/A	N/A
Growth	OFC 59.5 cm (48th) Height 87 cm (47th) Weight 13kg (62nd)	N/A	Birth weight 3.3 Kg	Weight 14.4 kg (<3rd) Length 108cm (25-50th) Microcephaly (<3 S.D)	Height 145cm Weight 33.9kg (41st)
Neurology					
Brain MRI	Normal	N/A	Normal	Cortical atrophy	N/A
Development	Walked 25 months	Walking : 18 months (bilateral vestibular areflexia)	Walking : 24 months (bilateral vestibular areflexia)	Unable to sit Absent speech Profound global delay	Normal
Intellectual disability	None	None	None	Severe	None
Autism	No	No	No	N/A	No
ADHD	No	No	No	N/A	No
Epilepsy	No	No	No	EEG abnormal	Seizure like episodes EEG abnormal
Motor	Hypotonia	Normal	Normal	Hypotonia	Hypotonia Myoclonus

					Poor exercise tolerance
Hearing	BLSNHL	BLSNHL	BLSNHL	BLSNHL	Normal
Ocular	Nystagmus	Hypermetropia, astigmatism, right amblyopia. Normal electroretinography	No	No	Nystagmus Photophobia
Cardiac Defect	No	No	No	No	Dilated cardiomyopathy
Gastrointestinal	No	No	No	Midgut malrotation Gastrostomy fed	Total parenteral nutrition
Respiratory	No	No	No	Choanal atresia	No

Supplementary table 1. *SLC12* gene family members and human disease.

Gene	Expression pattern	Disease	Phenotype
SLC12A1	Renal medulla Renal cortex	Bartter syndrome (OMIM 601678)	Hypokalaemia Hypotension
SLC12A2	Brain, Renal, GIT, vasculature, lung, glandular tissue, skin	Neurodevelopmental disorder/deafness	
SLC12A3	Renal medulla Renal cortex	Gitelmann syndrome (OMIM 263800)	Hypokalaemia Alkalosis
SLC12A4	Brain, Renal, GIT, vasculature, lung, glandular tissue	No known disease	
SLC12A5	Brain only	Epileptic encephalopathy (OMIM 606726)	Epilepsy
SLC12A6	Brain, Renal, GIT, vasculature, lung, glandular tissue	Andermann syndrome (OMIM 21800)	Neuropathy, agenesis of corpus callosum, intellectual disability
SLC12A7	Brain, Renal, GIT, vasculature, lung, glandular tissue	No known disease	
SLC12A8	Predominantly thyroid	No known disease	
SLC12A9	Predominantly spleen	No known disease	

Supplementary table 2. Additional *de novo* variants in patients with *SLC12A2* variants.

Case	Additional *de novo* variants	Pathogenicity
DDDP109851	None	
DDDP112246	None	
DDDP117398	None	
DDDP127331	MYO5C (15-52587832-A-G) EXOC3L2 (19-45735104-C-T)	Both predicted benign
DDDP129577	None	
Case 6 (Bedoukian)	None 46, XYY on CGH	
Case 7 (Renieri)	None	
Case 8 (AF)	N/A (father of case 9)	
Case 9 (AF)	NPEPPS (17-45669435_45669437dup) OR2T8 (1-248084798-C-T)	Both predicted benign

Supplementary material

Supplementary Methods

Ascertainment of individuals with *SCL12A2* variants

Exome sequencing in the DDD study (Wright *et al.*, 2015) has UK Research Ethics Committee approval (10/H0305/83, granted by the Cambridge South REC, and GEN/284/12 granted by the Republic of Ireland REC). For trio exome sequencing in the DDD study, saliva samples were collected (Oragene DNA collection kits, DNA Genotek, Kanata, ON, Canada) and DNA extracted (QIAsymphony, Qiagen, Venlo, Netherlands). Exome sequencing was performed at the Wellcome Trust Sanger Institute with Agilent SureSelect 55MB Exome Plus with Illumina HiSeq to investigate single nucleotide variants (SNVs) and small insertion-deletions (indels) in coding regions of the genome. An automated variant pipeline was used as previously described (Wright *et al.*, 2015). Probands with protein altering SNVs in *SLC12A2* were identified in patients 1-5. Further individuals with *SLC12A2* SNVs were identified via Genematcher. For patients 8 and 9 trio exome sequencing was performed using Roche® NimbleGen Medexome commercial kit (SeqCap® EZ Human Exome Probes v3.0) on a Illumina® NextSeq instrument. Variants were validated by Sanger sequencing for cases 6-9.

In silico assessment of pathogenicity of *SCL12A2* variants

The effects of variants in *SLC12A2* were assessed using Combined Annotation Depletion Dependent score (CADD)(Rentzsch *et al.*, 2019), MutationTaster (Schwarz *et al.*, 2010), FATHMM (Shihab *et al.*, 2013)and PolyPhen-2 (Shihab *et al.*, 2013). The presence of *SLC12A2* variants in human populations without NDD was queried using gnomAD. The excess in occurrence of *de novo SLC12A2* protein-altering (nonsynonymous and loss-of-function) variants in DDD was assessed using DenovolyzeR, an open source R package (Ware *et al.*, 2015). We used Denovolyzer with default parameters, number of subjects 9856 and DNM expected rate 0.8. The obtained p-value was then corrected for multiple testing with Bonferroni accounting for 19618 genes. Spatial clustering of missense *SLC12A2* DNMs was assessed using Denovonear, a python script that calculates the probability that DNMs within a gene cluster together based on per-gene mutability rate through one million simulations weighted by the context trinucleotide rate. We used MuPiT to map DNMs onto protein structure (Niknafs *et al.*, 2013).

Transcriptomic study of *SLC12A2* expression in developing human brain

Variations in *SLC12A2* transcript levels in the human brain among different developmental stages and anatomical regions were investigated using RNA microarray data from the BrainSpan Atlas of the Developing Human Brain (Miller *et al.*, 2014). Single cell RNA (scRNA) sequencing data from the human brain at 16-18 weeks gestation was obtained from(Pollen *et al.*, 2015). NetworkAnalyst was used to explore transcriptomic differences in scRNA data between *SLC12A2* expressing and non-expressing cells (Xia *et al.*, 2015).

***In silico* study of expression of *SLC12A2* splice isoforms**

RNA-seq data was downloaded from BrainSpan as a .csv file. Exon expression is reported in normalised form and quantified as CPM. The ratio of exon 21 to mean expression of all other exons was calculated to quantify expression of exon 21 containing transcripts. An exon 21:(mean expression exons 1 - 28) ratio <1.0 is taken to indicate the presence of transcripts which do not contain exon 21. Long-read RNA-seq data from developing mouse cochlea was queried using the internet portal (https://morlscrnaseq.org/). Evidence of alternative splicing of exon 21 was sought by visualising Sashimi plots (Ranum *et al.*, 2019).

***In vitro* assessment of NKCC1 ion transporter function**

NKCC1 function was assessed through K^+ influx measurements into *Xenopus laevis* oocytes injected with wild-type or mutant NKCC1 cRNA (Delpire *et al.*, 2011). All procedures involving animals were approved by Vanderbilt University Medical Center IACUC. Single oocytes were isolated from *Xenopus laevis* ovarian lobes using collagenase treatment (4 x 1.5 hr incubation with 10 mg/ml Collagenase D) and incubated overnight in L15 medium (250 ml Leibovitz medium + 200 ml water + 752 mg HEPES, pH 7.4, + 44 ug/ml gentamicin) at 16ºC. The following, day, groups of 20-25 oocytes were injected with 50 nl water containing 15 ng NKCC1 cRNA and returned to 16ºC for 3 days. Unidirectional K^+ influx was measured using Rb-86 isotope as tracer. Oocytes were first washed in a solution containing in mM: 96 NaCl, 4 KCl, 1 Ca_2Cl, 0.8 $MgSO_4$, 5 HEPES, pH 7.4, 200 mOsM, then preincubated for 15 min in identical solution containing 1 mM ouabain. The preincubation was then aspirated and oocytes were incubated in same ouabain-containing solution (or same solution made hypertonic with addition of 65 mM sucrose) with 5 uCi/ml Rb-86 for 1 hour. After 4 rapid washed with ice-cold solution, individual oocytes were placed in vials and lyzed with 200 ul 0.25N NaOH for 1 h, followed by 100 ul glacial acetic acid and 5 ml scintillation liquid. Aliquots of 5 ul of flux solutions were also counted to translate cpm into pmoles K^+. K^+ influx was expressed in pmoles/oocyte/hr.

Gene co-expression network analysis

From the scRNA data obtained from (Pollen *et al.*, 2015) we dichotomised cells into those robustly expressing *SLC12A2* and those with a count of 0 for *SLC12A2*. We then used NetworkAnalyst's "Gene Expression Table" function to upload normalised scRNA count data from *SLC12A2* expressing and non-expressing cells. Data was filtered with a variance filter of 15 and to remove low abundance transcripts. Data was not normalised as the data from Pollen et al was normalised prior to publication. Differential gene expression analysis was then performed (p<0.05, FDR 5%). This identified 589 differentially expressed genes. A gene co-expression network was then created, using the tissue specific setting (brain cortex). The network was created using the minimum setting, creating a network of 947 notes and 332 seeds. The network was visualised with autolayout and the reduced overlap setting. The label propagation algorithm was then used to identify 5 modules within the network. Function explorer was then used to perform an enrichment analysis of the modules.

Supplementary Results

Case reports

Patient 1 (270188) is a 12 month old boy with multiple congenital anomalies (iris coloboma, ventricular septal defect and tracheo-oesophageal fistula), the first child of non-consanguineous parents. Mother had unilateral coloboma but no NDD. He was born at 37 weeks gestation and spent a period of time in a special care baby unit due to feeding difficulties. At 12 months of age his weight was 3 Kg (64th percentile), length 68cm (1st centile) and orbito-frontal circumference 40 cm (1st percentile). At 12 months he had not sat independently or spoken any words. He had bilateral sensorineural deafness. Cranial magnetic resonance imaging (MRI) demonstrated cortical dysplasia. Comparative genomic hybridisation (CGH) was normal. Trio exome sequencing identified a de novo, heterozygous missense variant in *SLC12A2* (g.127450305C>T, p.(Ala327Val)).

Patient 2 (DDD4K.02331) is a 9 year old boy with a NDD consisting of autism and mild intellectual disability. He was born at term. There was no clinical evidence of hearing impairment. CGH was normal. Cranial MRI was not performed. Trio exome sequencing demonstrated a de novo, missense variant in *SLC12A2* (g.127469897G>A
, p.(Arg410Gln)).

Patient 3 (278327) is a 15-year old girl with a NDD consisting of global developmental delay, autism, motor stereotypy and hypotonia. She had bilateral sensorineural hearing impairment. She was born at 40 weeks gestation with initial feeding difficulties associated with reflux. At 15-years old her height was 165 cm (65th centile) and orbitofrontal circumference 55 cm (46th centile). She first walked age 3 years and was non-verbal. Cranial MRI was not performed. CGH was normal. Trio exome sequencing identified a de novo, heterozygous nonsense variant in *SLC12A2* (g.127503511G>A, p.(Trp892*)).

Patient 4 (293333) is a 3 year and 4 months old girl with a NDD consisting of spastic paraparesis and delay of speech and gross motor development. She did not have any hearing impairment. She was born at 40 weeks of gestation, there were no neonatal complications. At 3 years old her orbitofrontal circumference was 51.8 cm (82nd centile). A cranial MRI scan was reported as normal. CGH was normal. Trio exome sequencing identified a de novo, heterozygous missense variant in *SLC12A2* (g.12746683A>T, p.(Asn376Ile)).

Patient 5 (296317) is a 6 year old girl with a NDD consisting of severe global developmental delay. She was born at 39 weeks gestation and spent a period of time in a special care baby unit with nasogastric feeding. Her neonatal hearing assessment using AOAE and AABR showed clear responses bilaterally. Subsequently a percutaneous feeding tube was fitted because of an unsafe swallow. She has muscle hypertonia of all four limbs. She had left hip dislocation corrected with open reduction and left femoral osteotomy. She has a thoracolumbar scoliosis. She does not walk independently and is non-verbal. Cranial MRI demonstrated agenesis of the corpus callosum. CGH was normal. Trio exome sequencing demonstrated a frameshift variant in *SLC12A2* (g.127420201dup, p.(His186fs16)).

Patient 6 is a 21-year old boy with a NDD consisting of autism and intellectual disability. He is the only child of unrelated and healthy parents. Parents experienced three spontaneous abortions in the first trimester of gestation. He was born at term and showed a normal perinatal clinical history. He showed an almost normal motor development with ambulation acquired at 13 months. At two years of age he was diagnosed with autism spectrum disorder. Language was acquired at 6 years of age. Hearing tests were all normal. At 21 years height was 174 cm (36th centile), weight 64 Kg (27th centile) and orbitofrontal circumference 59 cm (+2.7 SD). Karyotype, oligo-array-CGH and *FMR1* triplet expansion region analysis were all normal. Trio

exome sequencing demonstrated a de novo, missense variant in *SLC12A2* (g.127466845G>C, p.Ala379Leu).

Patient 7 is a 25-month old boy with hypotonia, tongue fasciculation and developmental delay. He is the 4th child of healthy, non-consanguineous parents. He was born at term and failed his newborn hearing test. He has bilateral sensorineural hearing impairment with cochlear implant placed at 1-year of age. Nystagmus and tongue fasciculation were initially present but resolved by 1-year of age. Formal ocular examination was normal. At 25-months height was 87 cm (47th centile), weight 13 Kg (62nd centile) and orbitofrontal circumference 49.5 cm (48th centile). He walked independently at 25-months of age. He had speech delay associated with hearing impairment. A SNP-array demonstrated 47, XYY. Cranial MRI was normal. Trio exome sequencing demonstrated a *de novo*, missense variant in *SLC12A2* (g.127512895G>A, p.(Glu980Lys)).

Family S1585 (patient 8 and 9). Father and son affected by BNSHL with bilateral congenital vestibular areflexia with dominant pattern of inheritance. Cognitive abilities were not tested. Patient II.1(case 9), a full-term baby, presented with a congenital bilateral profound NSHL, revealed by abnormal neonatal hearing screening. He is the only child of unrelated parents. Unilateral cochlear implant was provided at age of two. Cranial growth showed a posterior plagiocephaly. Age of walking was 24 months old. Vestibular assessment was difficult but rotational vestibular ocular reflex was absent, suggesting a bilateral vestibular areflexia. His father (case 8), a 44 year old man, presented with a bilateral profound hearing loss and a bilateral vestibular areflexia (confirmed by video head impulse test, videonystagmography and absence of vestibular evoked myogenic potentials). The father's age of walking was 18 months old, which suggested a congenital onset of vestibular areflexia. Bilateral hearing aids were fitted at the age of 3 and were still used, only providing environmental sounds perceptions. The father's communication was sign language. He obtained a training vocational certificate and worked as a car washer. Trio exome sequencing demonstrate a missense variant segregating in *SLC12A2* (g.127512802G>A, p.(Glu979Lys)). The father was the only affected out of 7 siblings and his parents did not suffer hearing loss.

NKCC1 Flux analyses results

As seen in Figure S2, alanine at position 327 which is located towards the middle of transmembrane domain 2 (TMD2), is highly conserved within NKCC1 but not conserved among the different members of the Slc12a transporters. Mutation into valine (as in patient 1) results in a significant reduction in transport activity under both basal and stimulated conditions (Figure 2). As seen in Figure S1, arginine at residue 410 is a positively charged residue located within TMD4, highly conserved among the different NKCC1s and among the functionally well characterized cation-chloride cotransporters. Its mutation into glutamine (as in patient 2) decreased K^+ influx under both basal and stimulated conditions, but only by ~30% (Figure 2). The mutation from patient 3 (p.Trp892*) was tested next. The residue is located within the cytosolic COOH-terminal tail of the cotransporter and is highly conserved among all NKCC1 proteins listed and also among the different cation-chloride cotransporters (Figure S2). Termination of translation at that position results in a complete loss of NKCC1 function (Figure 2). The fourth mutation involves another transmembrane domain (TMD3) residue: asparagine at position 376. The asparagine is highly conserved within the Na^+-dependent cotransporters, while it is a conserved threonine in the Na^+-independent K-Cl cotransporters (Figure S1). Mutation into an isoleucine results in a significant decrease in NKCC1 function. Interestingly, the function is completely inexistent under basal or isotonic conditions but reduced by half under stimulated or hypertonic conditions (Figure 2). Patient 5 carries a mutation that terminates translation of the cotransporter open reading frame early within the cytosolic N-terminus. The mutation causes a frameshift causing the appearance of 16 new amino acids followed by a stop codon p.(His186AlafsTer17). Irrespective of this addition, the protein if translated, terminates before the first transmembrane domain, and thus no transport function can be associated to this fragment. As seen in Figure S1, the next mutant is located within the much less conserved carboxyl-terminal tail of NKCC1. It is part of a stretch of 4-5 negative residues that are mostly conserved from shark to human. Its mutation into lysine (p.(Glu980Lys), patient 7) significantly affected the level of flux under isosmotic (basal) conditions, but the cotransporter could be maximally stimulated by hypertonicity (Figure 2). Our last mutant from patient 6, p.(Ala379Leu), involves a residue in TMD3 that is neighboring the asparagine (at 376) mutated in patient 4. As seen in Figure 2, mutation of this conserved residue causes severe reduction in NKCC1 function under both basal and stimulated conditions. The effect of the p.(Glu979Lys) variant was not tested. However, we predict it to be deleterious based upon *in silico* analyses and the fact that the adjacent p.(Glu980Lys) variant reduces flux.
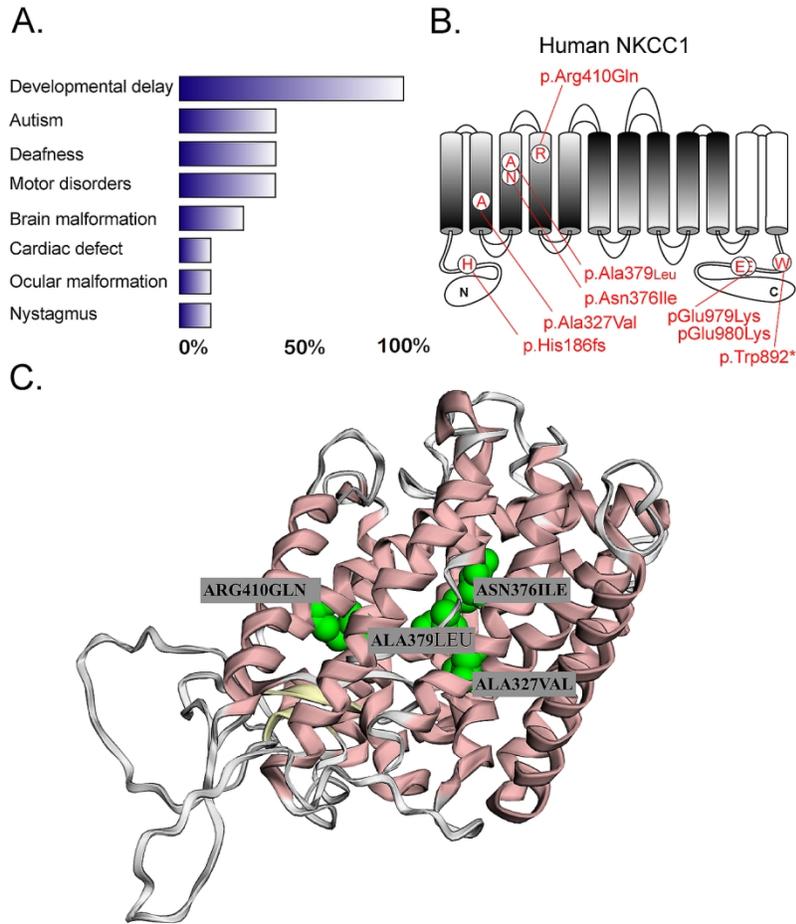
Figure 1. SLC12A2 missense variants.
A. Chart summarising frequency of clinical features in children with SLC12A2 de novo mutations and a NDD.
B. Schematic representation of NKCC1 with 12 transmembrane domains and inverted 5 TMs + 5 TMs
symmetry (shaded TMs), followed by TM11 and TM12 (white TMs). The position of patient mutations is
indicated in red. C. Three dimensional structure of NKCC1 demonstrating location of missense variants at
core of protein, with high potential to disrupt protein structure.

150x165mm (300 x 300 DPI)

Figure 2. Functional analysis of NKCC1 mutation in Xenopus laevis oocytes. A-B. Box plots demonstrating K+ influx measured in oocytes injected with water (negative control), water containing 15 ng wild-type (control) or mutant mouse NKCC1 cRNA. Horizontal line represents median, extent of box demonstrates interquartile range and whiskers minimum and maximum. K+ fluxes were measured under basal isosmotic (200 mOsM) or stimulated hyperosmotic (270 mOsM) conditions and are expressed in pmoles K+ per oocyte per hour. As the mutants were tested in different experiments on different oocytes, each mutant flux is accompanied by its own control flux. Statistical differences were determined using one way ANOVA. Each experiment represents 20-25 oocytes per mutation. Note that the mouse residue numbers are slightly different from human.

44x55mm (300 x 300 DPI)

Figure 3. Transcriptomic analysis of SCL12A2 expression in developing human brain. 3a. Box plots of microarray data comparing SLC12A2 expression in neuroanatomical regions of high and less active neurogenesis at 15 weeks gestation. Horizontal lines represent median, extent of box interquartile range and whiskers minimum and maximum. Expression is significantly higher in areas of active neurogenesis (*). 3b. Single cell RNA sequencing data demonstrating significantly higher expression of SLC12A2 in radial glia cells (RGC) compared to intermediate progenitor cells (IPC)(mann-whitney u-test, Z = -3.3, p = 0.001) and neurons (z = -5.2, p < 0.001). Units are counts per million reads (CPM).

209x297mm (300 x 300 DPI)
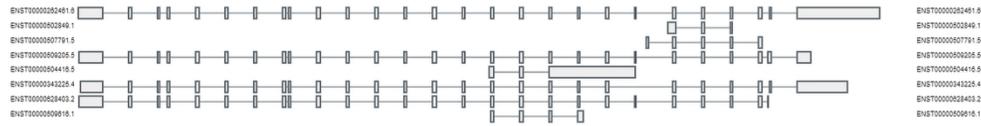
Figure S1. SLC12A2 splice isoforms.
Eight SLC12A2 isforms are identified by GTEx https://www.gtexportal.org/home/gene/SLC12A2. Note canonical isoform ENST00000262461.6 has 27 exons. ENST00000343225.4 has exon 21 spliced out and represents the minor isoform found in the developing brain.

201x29mm (300 x 300 DPI)

Figure S2. Sequence alignment of SLC12 gene family members.
The amino acid sequences of NKCC1 from eight species were aligned using VectorNti Suite 6.0 (Invitrogen). Portions representing the regions of interest were selected as follows: in A, a portion of TMD2 highlighting the conservation of the Ala327 residue; in B, a portion of TMD4 showing conservation of Arg410; in C, amino acids within a portion of the carboxyl-terminus showing conservation of Trp892; in D, portion of TM3 showing conservation of residues Asn376 and Ala377; and in E, a portion of the carboxyl-terminus showing some degree of conservation of residues Glu979 and Glu980. To assess conservation between the different members of the SLC12A gene products, the amino acids from human NKCC1, NKCC2, NCC, and KCC1-4 were also aligned. In F, residue Ala327 is less conserved among the different cotransporters; in G, Arg410 is highly conserved; in H, Trp758 is also a highly conserved residue; and in I, Asn376 and Ala377 are only conserved within the Na+-dependent cotransporters. Identical residues across are highlighted by yellow background, whereas conserved substitutions are shown in green or blue background. Non-conserved residues are shown in white background.

173x205mm (600 x 600 DPI)
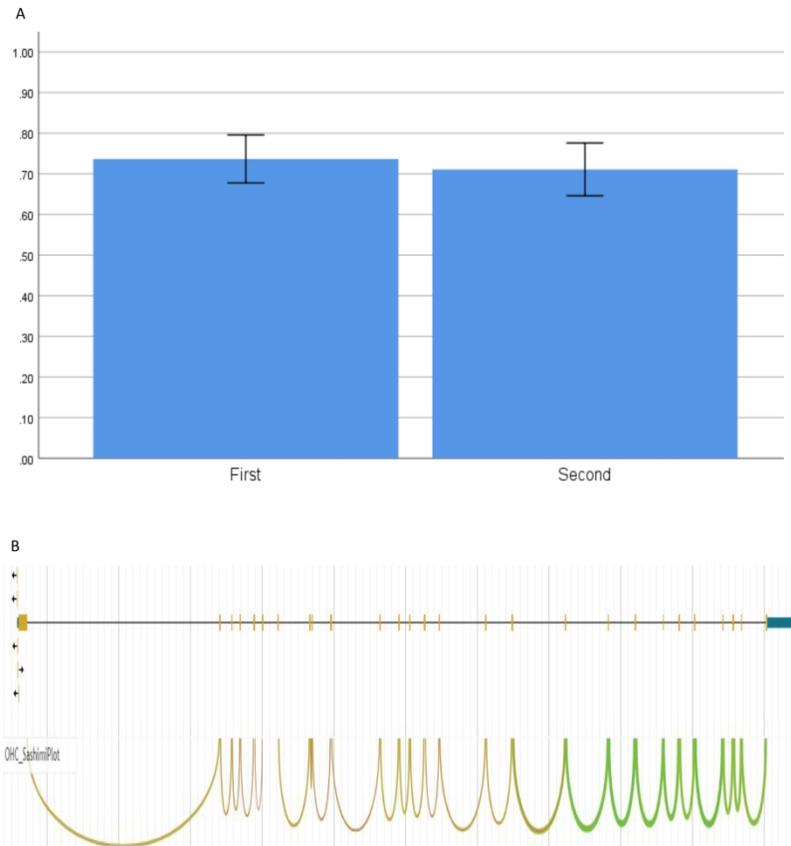
Figure S3. Expression of alternatively spliced exons and isoforms of SCL12A2.
a. The ratio of exon 21 expression to expression of all other 27 exons of SLC12A2 is <1.0 in the first and second trimester of pregnancy. This demonstrates the presence of transcripts which do not contain exon 21.
b. Image captured from https://morlscrnaseq.org/ an atlas of long-read single cell RNA expression data from murine inner ear. Sashimi plot demonstrates no alternative splicing events at exon 21. The only transcript expressed in mouse cochlea therefore contains exon 21.

209x297mm (300 x 300 DPI)

A

B

Reactome		GO Biological Process		GO Cellular Compartment	
				Pathway	FDR
Pathway	FDR	Pathway	FDR		
L1CAM interactions	0.000854	Axonogenesis	0.00237	Neuron projection	0.000166
Axon guidance	0.000854	Cell morphogenesis involved in differentiation	0.00237	Cytosol	0.000166
Translocation of GLUT4	0.00392	Neuron projection development	0.00331	Synapse part	0.00644
Apoptosis	0.00907	Cytoskeleton_dependent intracellular transport	0.00335	Synapse	0.0208
Glucose metabolism	0.00907	Neuron development	0.00364	Vesicle	0.0281
Transferrin endocytosis and recycling	0.0156	Interaction with host	0.00391	Cell junction	0.0295
Developmental Biology	0.0169	Axon guidance	0.0217	Cell projection	0.0295
NICD traffics to nucleus	0.0169	Neuron differentiation	0.0304	Protein complex	0.0303

Figure S4. Gene co-expression network analysis of single cell RNA data from SLC12A2 expression cells.
A. Figure shows module 1 identified by gene co-expression network based on differentially expressed genes created by NetWorkAnalyst.   The degree filter was used so that only hub genes are shown in this figure.
B. Table of significantly enriched terms for whole network.

210x297mm (600 x 600 DPI)