In Vitro Recapitulation of the Site-Specific Editing (to Wild-Type) of mutant *IDS* mRNA Transcripts, and Characterization of IDS Protein Translated from the Edited mRNAs

Susanna Lualdi,[1] Genny Del Zotto,[2#] Olga Zegarra-Moran†,[3#] Nicoletta Pedemonte,[3] Fabio Corsolini,[1] Maurizio Bruschi,[4] Valeria Tomati,[3] Giulia Amico,[1] Giovanni Candiano,[4] Andrea Dardis,[5] David N. Cooper[6] and Mirella Filocamo[1*]

[1]UOSD Centro di Diagnostica Genetica e Biochimica delle Malattie Metaboliche, Istituto Giannina Gaslini, Genova, Italy; [2]Core Facilities, Istituto Giannina Gaslini, Genova, Italy; [3]UOC Genetica Medica, Istituto Giannina Gaslini, Genova, Italy; [4]Laboratory on Physiopathology of Uremia, Istituto Giannina Gaslini, Genova, Italy; [5]Regional Coordinator Centre for Rare Diseases, University Hospital Santa Maria della Misericordia, Udine, Italy; [6]Institute of Medical Genetics, School of Medicine, Cardiff University, Heath Park, Cardiff CF14 4XN, UK

[*] **Corresponding author**:
e-mail: mirellafilocamo@gaslini.org

[#] These authors contributed equally to this work

†Deceased January 24, 2017

ABSTRACT

The transfer of genomic information into the primary RNA sequence can be altered by RNA editing. We have previously shown that genomic variants can be RNA-edited to wild-type. The presence of distinct 'edited' *IDS* mRNA-transcripts *ex vivo* evidenced the correction of a nonsense and frameshift variant, respectively, in three unrelated Hunter syndrome patients. This phenomenon was confirmed in various patient samples by a variety of techniques, and quantified by single nucleotide primer extension. Western blotting also confirmed the presence of IDS protein similar in size to the wild-type. Since preliminary experimental evidence suggested that the 'corrected' IDS proteins produced by the patients were similar in molecular weight and net charge to their wild-type counterparts, an *in vitro* system employing different cell types was established to recapitulate the site-specific editing of *IDS* RNA (U–to-C conversion and U deletion), and to confirm the findings previously observed *ex vivo* in the three patients. In addition, confocal microscopy and flow cytometry analyses demonstrated the expression and lysosomal localization in HEK293-cells of GFP-labeled-proteins translated from edited *IDS* mRNAs. Confocal high-content analysis of the two patients' cells expressing wild-type or mutated IDS confirmed lysosomal localization and showed no accumulation in the Golgi or early endosomes.

Key words: RNA editing conversion and deletion; mutation correction; edited IDS-transcripts to wild-type; confocal microscopy; imaging flow cytometry; western blot; SDS-PAGE; single nucleotide primer extension; expression vectors

Introduction

The transfer of genomic information into the primary RNA sequence can be altered by RNA editing at either the transcriptional or post-transcriptional level. Initially, this mechanism was thought to be quite rare in human cells [Bass, 2002]. However, over the past few years, multiple studies, both computational and experimental, have shown that RNA editing contributes extensively to the complexity evident at both the transcriptome and proteome levels [Li et al., 2011; Chen and Bundschuh, 2012; Kleinman et al., 2012; Peng et al., 2012; Zhu et al., 2012; Chen, 2013; Rieder et al., 2013; Bazak et al., 2014; Daniel et al., 2014; Li and Maso 2014; Savva and Reenan 2014; Xu and Zhang 2014]. In facilitating the genome-wide comparison of DNA and RNA sequences, the development of next generation sequencing (NGS) technology has served to greatly speed up studies of RNA editing [Piskol et al., 2013; Chen et al., 2014]. RNA editing is now recognized as a key level of regulation that is responsible for modulating a variety of cellular functions including protein activity, alternative mRNA splicing, and the alteration of miRNA target sites [Kim et al., 2016]. It should however be appreciated that the amount of RNA editing observed actually represents a population average of cells exhibiting a broad spectrum of rates of editing leading to the emergence of subsets of cells with differing informational content at the RNA level [Harjanto et al., 2016].

A number of different forms of mRNA editing have been reported in the human genome (dbRES) [http://bioinfo.au.tsinghua.edu.cn/dbRES] and several mechanisms have been described, ranging from co-, post- or co-post-transcriptional nucleotide modification, to nucleotide insertion or deletion [Bazak et al., 2014]. In human, RNA editing was first described as a C-to-U conversion in apolipoprotein B (*APOB*) primary transcripts by hydrolytic deamination giving rise to major alternative isoforms in the intestine (APOB48) and liver (APOB100), respectively [Wedekind et al., 2003; Navaratnam and Sarwar, 2006]. RNA editing is also known to modify RNAs in such a way as to introduce amino acid substitutions causing human genetic disease [Slotkin and Nishikura, 2013].

Cytidine (C) to uridine (U) and adenosine (A) to inosine (I) transitions are the best characterized mammalian mRNA editing mechanisms reported to date. Generally, in mammalian cells, cytidine to uridine (C to U) and A to I changes have been observed in both coding and noncoding sequences, including microRNAs; dysregulation can lead to pathogenesis [Gott, 2011; Maas et al., 2006; Grohmann et al., 2010; Ramaswami and Li, 2016].

Far less common is the U-to-C editing of mRNA (and tRNA) that has been observed in both the nucleus and mitochondria of mammals [Villegas et al., 2002; Niavarani et al., 2015]; at present, the underlying enzymatic processes have not been elucidated. In human, the transcript of the Wilms' tumor susceptibility gene, *WT1*, undergoes RNA editing by U to C conversion, thereby affecting the transcriptional repression function of the WTl protein [Sharma et al., 1994; Mrowka and Schedl, 2000]. U-to-C RNA editing has also been identified in long primary mi-RNAs (pri-miRNAs) [Blow et al., 2006; Griffiths-Jones et al., 2008].

An additional mechanism, U deletion mRNA editing, was first reported as a posttranscriptional modification of the RNA sequence in kinetoplastids, a group of flagellate protozoa [Estévez and Simpson, 1999; Osato et al., 2009]. However, evidence has emerged for the fairly frequent occurrence of U insertion editing in the human genome, where it represents a mechanism capable of creating different protein isoforms [Zougman et al., 2008]. Although the site specificity and basic mechanism of U insertion editing have not been experimentally demonstrated in all cases [Chen and Bundschuh, 2012], it has been shown in the trypanosomatids that specificity is determined by complementary guide RNAs (gRNAs) which act in *trans* to guide the RNA editing core complex (RECC, the 20S editosome) to the site of U-insertion or U-deletion by base pairing [Cruz-Reyes et al., 2001; Aphasizhev and Aphasizheva, 2011]. It was recently reported in trypanosomatids that this type of editing serves to correct frameshifts, introduce translation punctuation (start and stop) signals, and can even add hundreds of uridines to create novel protein-coding sequences [Aphasizheva and Aphasizhev, 2016].

More recently, a strategy based upon artificial RNA-guided editing machinery has been developed to re-program genetic information at the RNA level and it has been shown that such a strategy is capable of repairing disease-relevant genes, such as *CFTR*, not only in mammalian cell culture but also in simple organisms [Reautschnig et al., 2016]. Previously, we reported that human *IDS* gene variants affecting function were RNA-edited *ex vivo* by an as yet unidentified mechanism, which led to the expression of variable amounts of non-mutated (i.e. wild-type) *IDS* transcripts in a tissue-dependent fashion [Lualdi et al., 2010]. These mutational events were observed in three male patients affected by Hunter syndrome or mucopolysaccharidosis type II (MPS II, OMIM# 309900), a rare X-linked recessive lysosomal disorder caused by the deficiency of iduronate-2-sulfatase (IDS; EC 3.1.6.13) resulting from hemizygous genomic variants in the *IDS* gene (OMIM #309900; GenBank Accession No. NG_011900.1). Nearly 60% of the >550 different *IDS* variants reported to date (see the Human Gene Mutation Database) [http://www.hgmd.cf.ac.uk/ac/index.php; Stenson et al., 2017], are sequence variants with the potential to impact RNA processing in some way. This is how, during the course of *IDS* RNA analysis, we observed, in 3 unrelated male (hemizygous) Hunter syndrome patients, evidence of *ex vivo* correction (at the RNA level) of their gene lesions characterized at the DNA level [Lualdi et al., 2010]. The presence of two distinct "edited" *IDS* mRNA-transcripts (U-to-C and delU, correcting a nonsense variant and a frameshift variant, respectively) in different patients, was confirmed in various tissues and cell lines from the patients by several different techniques and quantified by single nucleotide primer extension (SNuPE). Western blotting analysis confirmed the presence of IDS precursor protein similar in size to the wild-type. We were left with the intriguing possibility that an mRNA correction mechanism akin to RNA editing had been activated in our patients. Here, we have confirmed by SDS-PAGE Western blot that the 'corrected' IDS-proteins, synthesized from edited *IDS* mRNAs produced by the patients, are similar in terms of molecular weight and net charge to their wild-type counterparts. Prompted by this experimental evidence, an *in vitro* system, employing different cell types, was used to recapitulate the site-specific editing of RNA (both U–to-C conversion and U deletion), and to confirm the findings previously observed *ex vivo* in

Field Code Changed

Formatted: English (United States)

the three patients. In addition, confocal microscopy, flow cytometry and imaging flow cytometry analyses provided evidence for both the expression and lysosomal localization in HEK293 cells of GFP-labeled proteins translated from edited *IDS* mRNAs. Confocal high-content analysis of the two patients' cells expressing wild-type or mutated IDS confirmed lysosomal localization and showed no accumulation in the Golgi or early endosomes.

MATERIALS AND METHODS

Patient and control samples

Fibroblast and lymphoblast cell lines were grown from the two unrelated Italian MPS II patients (pts) respectively hemizygous for the *IDS* genomic variants, c.22C>T and c.10insT (by contrast, at the RNA level, the presence of edited *IDS* mRNA transcripts c.22U>C and c.10delU, respectively, was evident). The c.22C>T and c.10insT variants, both of which are located in exon 1, predicted the truncating variants p.R8X and p.P4SfsX43, respectively. Fibroblast and lymphoblast cell line samples from a wild-type male individual were used as negative controls. Fibroblasts harbouring the *IDS* mutant p.W12X (c.35G>A), which proved not to be prone to RNA editing, were also included in the experiments.

The patient's samples were obtained from the "Cell Line and DNA Biobank from patients affected by Genetic Diseases" [Filocamo et al., 2014]. Owing to the paucity of Biobank samples, only one of two patients known to harbour the c.22C>T (R8X) was studied in the present work.

Ethical aspects. Following ethical guidelines, all samples for analysis and storage were obtained with the patients' (and/or the legal guardian's) written informed consent. Consent was sought using a form approved by the local Ethics Committee.

Variant nomenclature. According to current variant nomenclature recommendations (Human Genome Variation Society) [http://varnomen.hgvs.org/; den Dunnen et al., 2016], the above mentioned variants, p.R8X, p.P4SfsX43 and p.W12X, should be designated p.R8*, p.P4Sfs*43 and p.W12*, respectively. However, as the present study represents a continuation of that published previously [Lualdi et al., 2010], the former variant designation is used throughout the text in order to maintain continuity.

Field Code Changed

Cell culture

In-house established human fibroblasts and the commercially available human cell lines, HeLa and HEK293, were cultured according to standard procedures and maintained in RPMI medium (EuroClone, Gibco, Paisley, UK) containing 15% FCS and penicillin/streptomycin, in a humidified atmosphere containing 5% CO_2 at 37°C. The human fibroblasts were tested for mycoplasma contamination before use.

Biochemical studies on IDS protein from patient cell lines

Cells collected for all electrophoresis experiments were counted, washed in TBS and solubilised in Laemmli sample buffer for SDS-reduced-PAGE and 2D-PAGE, or in Tris/Boric acid/ EDTA buffer for Native-PAGE. Sample concentrations were determined by means of a bicinchoninic acid protein assay.

***Native-PAGE*.** Samples (50 µg) were solubilised in 90 mM Tris, 80 mM Boric acid and EDTA buffer at pH 8.6 containing phosphatase and protease inhibitor cocktail and loaded onto a gel (T% 8–16) with the same buffer in a Protean II XI system (Bio-Rad, Hercules, CA, USA).

SDS-reduced-PAGE. Samples (50 µg) were solubilised in Laemmli sample buffer containing 50 mM DTE and phosphatase and protease inhibitor cocktail. Samples were then boiled for 10 min and separated in a SDS-PAGE (T% 8–16) with a Protean II XI system (Bio-Rad).

2D-PAGE. The solubilised samples (200 µg) were cleaned and precipitated using an ice-cool mixture of tri-n-butyl-phosphate:acetone:methanol with ratio of 1:12:1 (TBPAM) as previously reported [Bruschi et al., 2005] with some modifications. Briefly, this mixture was added to each sample to reach a final acetone concentration of 80% (v/v) and incubated for 90 min at 4°C. The precipitate was pelleted by centrifugation at 2800 x g for 20 min at 4°C. Supernatant was discarded and the pellet was washed with 1 ml tetrahydrofuran (THF) precooled on ice. After centrifugation at 10000 × g for 10 min at 4°C, the pellet was air dried, dissolved in the sample focusing solution (7M urea, 2M thiourea, 4% (w/v) CHAPS, and 50 mM DTE) and loaded onto home-made non-linear pH 3–10 soft-IPG strips (18 cm long, 3 mm wide, 0.5 mm thick) [Bruschi et al., 2003]. Strip re-swelling was carried out overnight at room temperature in the focusing solution, 7M urea, 2M thiourea, 2% (w/v) CHAPS, 15 mM DTE, and 0.6% (v/v) carrier ampholyte cocktail, containing 40% of the pH 3.5–10 and 60% of the pH 4–8 intervals.

The proteins were focused at ≤ 50 mA per strip at 20°C, using a progressively increasing voltage for a total of 90000 Vh. An equilibration step was performed for 30 min in the equilibration buffer containing 6 M Urea, 50 mM Tris-HCl pH 8.8, 2% (w/v) SDS, 30% (v/v) glycerol, and a trace of bromophenol blue. In the second dimension, the proteins were separated on an SDS-PAGE gel (T% 8–16) with a Protean II XI system (Bio-Rad).

Western Blotting. After the electrophoretic separation, the samples were transferred onto nitrocellulose membrane (Protean BA, Schleicher & Schuell, Dassel, Germany) with a Novablot semi-dry system, 3% w/v BSA saturated in TBS and incubated separately with 1:1000 anti-IDS (Santa Cruz Biotechnologies, CA, USA), anti-IDS (Abnova, Taiwan, Taipei), anti-α-tubulin (Sigma, Saint-Louis, USA) in 3% w/v BSA in TBS-Tween 0.05% v/v (TBS-T). Membranes were then rinsed in

TBS-T and incubated with secondary antibody. Chemiluminescence was used for immunodetection. Images were digitalized by mean of VersaDoc 4000 (Bio-Rad) and analyzed with QuantityOne software (Bio-Rad).

Immunoprecipitation. Immunoprecipitation by anti-IDS antibody (Abnova, Taiwan, Taipei) was performed with 1 mg protein from whole cell lysates. After pre-cleaning by protein A Sepharose, the samples were incubated with anti-IDS antibody overnight at 4°C with gentle rotation. The protein A-Sepharose beads were then washed extensively with washing buffer and eluted with 2D-PAGE focusing solution. The bound proteins eluted from the beads were separated on 2D-PAGE and the gels were subjected to silver staining.

Construction of full-length *IDS* cDNA vector for *in vitro* expression

RT-PCR products of full-length *IDS* cDNAs, obtained using primers located in the 5'UTR and 3'UTR respectively [Lualdi et al., 2010], were cloned into the plasmid vector pcDNA3.1/V5-His–TOPO, containing the CMV promoter and T7 promoter (see map) [http://tools.invitrogen.com/content/sfs/manuals/pcdna3.1topota_man.pdf], to generate *IDS* cDNA constructs as follows: (i) mutants p.R8X (c.22C>T) and p.P4SfsX43 (c.10insT) both of which proved to be RNA editing-prone; (ii) wild-type (negative control); (iii) mutant p.W12X (c.35G>A) (positive control) which shares with the two variants under investigation the characteristics of (a) affecting the same *IDS* gene region (signal peptide), and (b) introducing a premature termination codon, but which differs from these aforementioned variants in terms of not being RNA editing-prone.

Field Code Changed

The four different constructs were transformed into competent *E. coli* cells using the One Shot TOP10 Chemical Transformation Kit (Invitrogen). Clones were isolated by QIAprep Kit (Qiagen). The presence of the insert in the construct was confirmed by restriction enzyme digestion, whilst the sequences of all constructs were checked for potential additional variants by DNA sequence analysis.

Transfection of the full-length *IDS* cDNA constructs and characterization of transcripts by Single Nucleotide Primer Extension (SNuPE) analysis

Eight µg of each vector were transfected by Lipofectamine® Reagent (Invitrogen), without antibiotics according to the manufacturer's recommended instructions, into the different cell types (>95% confluent): viz. fibroblast cell lines from both patients (p.R8X and p.P4fsX43) and HEK293 and HeLa cell lines (normal control). After 48 h/72 h, the cells were trypsinized and washed in PBS. Total cytoplasmic RNA was extracted using RNeasy Mini and Plus Micro kit (Qiagen) to prevent genomic DNA contamination, and reverse transcribed using the SuperScript® VILO™ cDNA Synthesis Kit (Invitrogen). To amplify the fragment containing all three variants under investigation, [c.22C>T (p.R8X)], [c.10insT (p.P4 SfsX43)] and [c.35G>A (p.W12X)], RT-PCR was performed with a pair of primers, the forward primer (5'-GGAGACCCAAGCTGGCTAGT-3') being complementary to the T7 promoter site (upstream of the *IDS* 5'UTR) and the reverse primer (5'-GGTCACATAGCCATTCTCCTTG-3') being located within exon 3. The use of the T7 primer allowed us to selectively amplify, and hence distinguish the exogenous *IDS* transcripts from the endogenous *IDS* transcripts. The amplified products were analyzed by electrophoresis on 1.5% agarose gels followed by staining with Atlas ClearSight DNA Stain (Bioatlas).

To quantify the relative amounts of edited/non-edited transcripts, SNuPE analysis of RT-PCR products was performed using specific primers for [c.22C>T (p.R8X)], [c.10insT (p.P4SfsX43)] and [c.35G>A (p.W12X)] according to the previously described procedure [Lualdi et al., 2010].

Construction of pAcGFP1-N vector containing full-length *IDS* cDNAs

RT-PCR products of full-length *IDS* cDNA were then cloned into the plasmid vector pAcGFP1-N In-Fusion Ready Vector (see map) [http://www.clontech.com/xxclt_ibcGetAttachment.jsp?cItemId=17882] according to the manufacturer's instructions, using the forward 5'-

Field Code Changed

AAGGCCTCTGTCGACGCCACCTGCTGCAGCCTGTCC-3' and the reverse 5'-AGAATTCGCAAGCTTAGGCATCAACAACTGGAAAAG-3' primers, respectively. The *IDS* cDNA sequence was added in-frame to that of the AcGFP1 coding sequence to generate an in-frame fusion protein to the N-terminus of AcGFP1. The resulting constructs contained: (i) normal WT-IDS (negative control); (ii) R8X(c.22C>T)-IDS mutant prone to the RNA editing-like mechanism, and (iii) the W12X(c.35G>A)-IDS mutant not prone to the RNA editing-like mechanism. The pAcGFP1-N vector harbouring IDS(p.P4Sfs) was excluded from the analysis because of its low protein expression as revealed by Western blot analysis.

The four different constructs were then transformed into Stellar competent *E. coli* cells by means of the In-Fusion HD Cloning Kit (Clontech). Clones were isolated by QIAprep Kit (Qiagen). All constructs were checked for accidentally-introduced variants by enzyme restriction digestion and sequencing analysis.

Confocal microscopy

For live-cell imaging experiments, HEK293 cells were cultured on glass coverslips. When 95% confluence was reached, cells were transiently transfected with 1.5 μg of either p-IDS(WT)-AcGFP1-N, p-IDS(R8X)-AcGFP1-N or p-IDS(W12X)-AcGFP1-N plasmids using Lipofectamine® Reagent (Invitrogen), as described above. After 48 h transfection, the nuclei of HEK293 cells were stained by adding 1 μg/ml Hoechst 33342 dye (Enzo Life Sciences) to the culture medium. After 15 mins, the medium was removed, cells were washed with phosphate-buffered saline (PBS) and then incubated for 30 mins in PBS containing 100 nM LysoTracker Deep Red (Invitrogen). LysoTracker is a weakly basic fluorescent dye that stains acidic organelles, in particular lysosomes. Cells were maintained for 30 mins at 37°C in a 5% CO_2-containing incubator. For confocal microscopy, the coverslip with stained cells was removed and placed, inverted, on a microscope slide. The borders were sealed to avoid evaporation of the thin liquid film sandwiched between the two glasses. For confocal microscopy, the cells were placed on the stage of a laser-scanning microscope (Leica SP8) under a

40X, oil immersion objective (N.A. 1.3). GFP was laser excited at 488 nm and the emission wavelength was collected between 495 and 590 nm. LysoTracker Deep Red was excited at 633 nm and the emission wavelength captured between 640 and 750 nm. Hoechst 33342 was excited at 405 nm and the emission captured between 410 and 460 nm. Excitation of the different fluorochromes was performed sequentially. After acquisition of each set of fluorescent images, a differential interference contrast (DIC) image was obtained to make clear the cells' position and morphology. Fluorescent image reconstruction and merging were performed using ImageJ software.

Transfection with vector p-IDS-AcGFP1-N in the HEK293 cell line for flow cytometric and imaging flow cytometry analyses

HEK293 cells, cultured on T25 flasks, were transiently transfected (>95% confluent) with 3.75 μg p-IDS(WT)-AcGFP1-N, p-IDS(R8X)-AcGFP1-N or p-IDS(W12X)-AcGFP1-N plasmids using Lipofectamine®Reagent (Invitrogen) without antibiotics according to the manufacturer's instructions. After 48 h, the cells were trypsinized and washed in PBS. HEK293 cells were stained with 100 nM solution of LysoTracker Deep Red (Invitrogen) in PBS for 30 mins at 37°C in 5% CO_2 for Imaging flow Cytometry analysis.

Flow cytometric analysis.

All flow cytometric samples were acquired on a four laser Gallios (Beckman Coulter) while, after acquisition, compensation and data analysis were performed with Kaluza software (Beckman Coulter). HEK293 cells were evaluated on the basis of their level of GFP protein expression.

Imaging flow cytometry (IFC) analysis.

Cells were analyzed in the ImageStreamX Mark II imaging flow cytometer (Amnis Corporation, Seattle, WA) equipped with a MultiMag system. Data were acquired using INSPIRE acquisition

software (Amnis, Merck). 5000 events were recorded using a 20x objective in order to perform quantitative analyses of samples whilst 1000 events where collected with both 40x and 60x objective lenses to obtain an improved view of the analyzed co-localization of GFP and LysoTracker.

Upon focus, cells were initially identified from the bright-field images by using an algorithm recommended by the manufacturer, the Gradient RMS (root mean square) feature. Since HEK293 is an adherent cell line that tends to form aggregates even when resuspended strongly, we took advantage of two other features of INSPIRE software, Area and Aspect Ratio (the smallest width divided by the longest; for single cells, this ratio should usually be 0.8-1), which allowed us to specifically focus our attention on single cell events. Thanks to these features, we were able to optimize sample collection by recording only those events corresponding to perfectly in-focus single cells. To perform this analysis, four ImageStream channels were used: channel 1 for the bright-field, channel 2 for GFP (excited by a 488nm laser), channel 5 for LysoTracker (excited by a 642nm laser).

Data were then analyzed using IDEAS 6.0.3 software (Amnis, Merck). Compensation was set on the basis of a matrix calculated by IDEAS software using single-stained samples (GFP and LysoTracker respectively) obtained with the same laser setting used for acquisition but with the bright-field light switched off, as suggested by the manufacturer

Confocal high-content imaging and analysis of IDS protein expression and subcellular localization. Fibroblasts from healthy or Hunter syndrome individuals were plated on clear-bottom black 96 well/plate suitable for high-content imaging. After 72 hours, cells were fixed with neutral-buffered formalin and permeabilized with Triton X-100 (0.2%). Fixed cells were then incubated with anti-IDS (R&D), anti-LAMP1 (Santa-Cruz), anti-giantin (Covance) and anti-EEA1 (GeneTex) antibodies at a concentration of 0.25 µg/ml for 2 h at 37°C. After washing with PBS, cells were incubated with AlexaFluor 647 anti-mouse IgG2b (for IDS), AlexaFluor 488 anti-mouse IgG1 (for LAMP-1) and AlexaFluor 546 anti-rabbit IgG (for giantin or EEA1) secondary antibodies (Life Technologies).

Confocal high-content imaging and data analysis of cells were performed using an Opera Phenix (PerkinElmer) equipped with two cameras and synchrony optics that allow simultaneous acquisition (without crosstalk) of two fluorophores. Cells were imaged using a 40X water immersion objective (N.A. 1.1). AlexaFluor 488 signals were laser excited at 488 nm and the emission wavelength were collected between 500 and 550 nm. AlexaFluor 546 was excited at 561 nm and the emission wavelength captured between 570 and 630 nm. Hoechst 33342 was excited at 405 nm and the emission wavelength captured between 435 and 480 nm.

Data analysis of co-localization was performed on approximately 100 cells per condition, by using the Harmony software (ver 4.5) of the Opera Phenix high-content system.

RESULTS

Biochemical studies on IDS protein from patient cell lines.

Western blots of SDS-reduced-PAGE and native-PAGE gels revealed a unique protein product, with a molecular weight comparable to that of IDS (MW=61.9kDa) in both patient samples (p.R8X and p.P4SfsX43); a reduced protein concentration was noted exclusively in the patient harboring p.P4Sfs, as compared to the control (Figure 1).

To verify the absence of any difference in molecular weight and/or net charge modification of IDS protein from the patients as compared to the control, a western blot of two-dimensional gel electrophoresis was performed. Unfortunately, the IDS antibodies employed did not detect any protein spots (data not shown). To resolve this problem, immunoprecipitation of the protein was performed by anti-IDS; as shown in Figure 2, at the appropriate molecular weight and isoelectric point of IDS protein (pI=5.45), a family of three protein spots was noted which exhibited similar expression to the IDS profile in western blot experiments.

Correction of the *IDS* gene lesions by editing at the RNA level in a transient expression study.

In an attempt to reproduce *in vitro* the previously documented RNA editing-like mechanisms [conversion of U to C in the case of c.22C>T (p.R8X) and a U deletion in the case of c.10insT (p.P4SfsX43)], we set up an expression system in different cellular models including human fibroblasts, HeLa and HEK293 cell lines. Specifically, the wild-type *IDS* cDNA and the mutant *IDS* cDNAs [c.22C>T (p.R8X), c.10insT (p.P4SfsX43) and c.35G>A (p.W12X)] were inserted into the plasmid vector pcDNA3.1/V5-His–TOPO, containing the CMV promoter and T7 promoter, and then introduced by lipid-mediated transfection into: (i) fibroblasts from the two patients exhibiting the RNA editing-like mechanism(s), conversion and deletion, respectively; (ii) normal fibroblasts; and (iii) commercially available human cell lines viz. HeLa and HEK293. To quantify the amount of nucleotide correction (U>C and delU, respectively), SNuPE of RT-PCR products was performed with specific primers as previously described [Lualdi et al., 2010].

The SNuPE results showed that the plasmid constructs harbouring p.R8X and p.P4SfsX43 yielded the expected transcripts (i.e. both mutated and corrected to wild-type) *in cellula*, not only when they were transfected into Hunter syndrome patient fibroblasts (which of course display the RNA editing-like phenomenon) but also when they were transfected into healthy donor fibroblasts and commercially available cell lines viz. HeLa and HEK293. As indicated in Table 1, the extent of nucleotide correction ranged from 17% to 53% for the U deletion [c.10insT (p.P4SfsX43)], and from 0% to 63% for the conversion of U to C [c.22C>T (p.R8X)], depending upon the type of transfected cell. In agreement with the *ex vivo* results observed in the patients in the previous study [Lualdi et al., 2010], no correction occurred *in cellula* with c.35G>A (p.W12X). Consistent with this finding, capillary electrophoresis revealed a bimodal pattern indicating the presence of two distinct populations of *IDS* mRNA transcripts (Figures 3 A-B) when the mutant constructs, c.10insT (p.P4SfsX43) and c.22C>T (p.R8X), were transfected into the (same) patient's fibroblasts [i.e. c.10insT (p.P4SfsX43) into patient fibroblasts harbouring c.22C>T (p.R8X) and *vice versa*] or into a

control. By contrast, Figure 3C depicts a unimodal capillary electrophoretic pattern when the mutant c.35G>A (p.W12X) construct (positive control) was transfected into all cell types, a pattern that indicates the presence of 100% *IDS* mutated (i.e. unedited) transcripts. The results obtained suggest that the proportion of edited *IDS* mRNA transcripts varied depending upon the nature of the variant involved and the cell type in which the constructs were expressed.

Confocal microscopy analysis: evidence of IDS-GFP expression and IDS-GFP lysosomal localization in HEK293 cells transfected with the pAcGFP1-N vector harbouring IDS(WT), IDS(R8X) and IDS(W12X).

Evidence of expression. HEK293 cells transfected with both p-IDS(WT)-AcGFP1-N and p-IDS(R8X)-AcGFP1-N vectors showed abundant green GFP staining, indicating the high level expression of chimeric protein which was distributed within the cytoplasm of the cell with an evident colocalisation in the lysosomes (Figure 4 A-B). By contrast, expression levels of the IDS(W12X)-GFP fusion protein, bearing a stop codon that is predicted to lead to premature termination of translation (but empirically shown not to be prone to mRNA correction) were, as expected, scarcely visible (Figure 4 C).

Evidence of lysosomal localization. LysoTracker stained acidic cellular compartments in all cells. Overlaying red and green fluorescence into a merged image revealed those regions where both IDS(WT)-GFP and IDS(R8X)-GFP co-localized with Lysotracker (yellow staining indicated by arrows), suggesting IDS-GFP lysosomal localization, albeit only partial (Figure 4, fourth column, A and B).

Flow cytometry (FC) analysis: quantification of IDS-GFP expression in HEK293 cells transfected with the pAcGFP1-N vector carrying IDS(WT), IDS(R8X) and IDS(W12X).

As an adherent cell line, HEK293 cells tend to stick together. For this reason, and in order to avoid doublets or triplets that could alter the proportions of positive GFP events, Side Scatter (SSC) peak

width was plotted against SSC peak height, allowing clear discrimination between single cells and doublets (Figure 5 A). These doublets were then gated out and single cells were analysed for their GFP expression. As shown in Figure 5 B, the samples expressing IDS(WT)-GFP and IDS(R8X)-GFP fusion proteins showed comparable proportions of positive events with similar mean fluorescence intensity (MFI). By contrast, the samples expressing IDS(W12X)-GFP showed only slight positivity, being present in only 16% of the cell population, with a noticeably lower MFI. Flow cytometry quantification analysis therefore confirmed that the amount of corrected-fusion-IDS protein was comparable to that of the wild-type IDS protein.

Imaging Flow Cytometry (IFC) analysis: evidence of lysosomal localization and quantification of IDS-GFP expression in HEK293 transfected cells with pAcGFP1-N vector harbouring IDS(WT) and IDS(R8X).

In order to perform a semi-quantitative analysis of the samples from transfected cells, we acquired 5000 cellular events using a 20x objective. We further optimized the analysis by focusing exclusively on events that corresponded to perfectly in-focus single cells (Figure 6 A-B). Data obtained for both IDS(WT)-GFP and IDS(R8X)-GFP fusion protein expression in HEK293 showed comparable percentages (Figure 6 C), thereby confirming what had already been observed by flow cytometry. In contrast to classical cytometry, through which we could only establish if two markers of interest (LysoTracker and GFP) were coexpressed in the same cell, IFC also allowed us to determine whether or not LysoTracker and GFP co-localized. To identify the cell sub-population bearing the co-localization of interest, we applied a feature called “bright detail similarity” to a custom mask (Figure 6 D). Through the algorithm obtained, we were able to define a cluster of the best images of those events exhibiting high expression of both markers in the same position of the cell surface at the same time. This analytical strategy was also applied to cellular events acquired with both 40x and 60x objectives, in order to perform a qualitative analysis leading to good lysosomal localization pictures of both IDS(WT)-GFP and IDS(R8X)-GFP fusion proteins expressed *in vitro* (Figure 6 E-F). Overall,

analysis of the data allowed us to confirm that the IDS protein produced using the edited mRNA as template was comparable to the wild-type IDS protein both in terms of its amount and its lysosomal localization.

Confocal high-content screening system: imaging and analysis of IDS protein expression and subcellular localization in patient cell lines. High-content imaging analysis was performed on fibroblasts from control and Hunter syndrome individuals using the Opera Phenix system (PerkinElmer). This analysis revealed that in control fibroblasts wild-type IDS protein is expressed and co-localizes with LAMP-1, a lysosomal marker. Co-localization of the two signals was higher than 90% (Figure 7A). By contrast, the expression of W12X-IDS in fibroblasts from a Hunter syndrome patient was extremely low (an ~100-fold decrease as compared to the expression of wild-type IDS) with no co-localization with the lysosome marker LAMP-1 (Figure 7B). However, fibroblasts from Hunter syndrome patients bearing either the p.R8X (Figure 7C) or p.P4Sfs (Figure 7D) mutation display very high levels of expression of the edited-to-wild-type mutant protein (a ~50-fold increase as compared to expression of wild-type IDS) that accumulates inside the cells. Co-localization of edited-to-wild-type R8X- or P4Sfs-IDS with LAMP-1 was low (approximately 10%). We analyzed also the co-localization of mutant IDS with the Golgi marker giantin or the early endosome marker EEA1, and the results demonstrated that edited-mutant IDS did not accumulate in the Golgi nor in early endosomes (Figure 7C-D).

Review of available literature and bioinformatics analysis.

To ascertain whether there might be other human *IDS* target sites for RNA editing in addition to those investigated here, we comprehensively screened the relevant literature: a computational study that aimed to determine the extent of RNA editing in the human transcriptome identified a single U-to-A editing site in an intronic position of the unspliced *IDS* mRNA (chrX 1,48E+08) [Kleinman et al., 2012]. We also surveyed bioinformatics software resources available online. The RADAR database [http://rnaedit.com/] lists three RNA A-to-I editing events affecting 3 distinct sites in the *IDS* gene, 2

Field Code Changed

intronic (chrX 1,48567508E+08 and 1,48567509E+08) and 1 exonic (chrX 1,48577934E+08). We also queried the DARNED database [http://darned.ucc.ie/] but no *IDS* editing sites were found.

DISCUSSION

The twin aims of this work were (i) to elucidate further the biochemical features of the IDS proteins encoded by the mutant mRNAs (p.P4SfsX43 and p.R8X) edited-to-wild-type, which we previously observed in three patients affected by Hunter syndrome and (ii) to recapitulate *in vitro* the site-specific editing of RNA, viz. the U–to-C conversion and U deletion [Lualdi et al., 2010].

The biochemical analysis panel, including the western blotting of proteins in native-PAGE, reducing SDS-PAGE and two-dimensional gel electrophoresis of immunoprecipitated proteins, confirmed the presence, in the patient's cell line harbouring p.R8X, of wild-type IDS protein with similar amount, molecular weight and charge properties to those noted in the control. In the patient's cell line harbouring p.P4Sfs, the wild-type IDS protein exhibited similar molecular weight and charge properties to that of the control, but a comparatively low amount of protein compared to the control. Therefore, the wild-type protein detected could only have been translated from the edited *IDS* transcripts identified *ex vivo* in the two Hunter syndrome patients. The *in vitro* expression system, developed to investigate these observations further, employed different human cell lines and successfully recapitulated the site-specific editing of mutated RNA.

Intriguingly, *in cellula*, plasmid constructs harbouring both p.R8X and p.P4SfsX43 yielded both mutated and edited transcripts, not only when they were transfected into fibroblasts cultured from the same Hunter syndrome patients (which were known to be positive for the RNA editing phenomenon) but also when they were transfected into healthy donors' fibroblasts and commercially available HeLa and HEK293 cell lines (Table 1). Whether editing occurs or not therefore appears to be an inherent property of the target sequence. However, the observed extent of the nucleotide-correction may depend upon the type of transfected cells but may also be regulated by RNA secondary structure and the temporal expression of RNA editing enzymes in the same cell population, as described for known

editing mechanisms, such as that occurring in APOB [Hwang et al., 2016]. Taken together, these results demonstrated that our *in vitro* system was able to recapitulate the putative correction mechanism and that this mechanism, acting independently of cell type, was likely to be dependent upon intrinsic characteristics of the editing sites.

In addition, at the protein level, confocal microscopy yielded evidence both for IDS-GFP expression and IDS-GFP lysosomal localization, albeit only partial, in HEK293 cells transfected with the pAcGFP1-N vector carrying IDS(p.R8X) (Figures 4B). Similar findings were obtained by flow cytometry and imaging flow cytometry analyses (Figures 5 and 6). Based on these results, it is clear that our *in vitro* system was able to recapitulate the observed mechanism *ex vivo*, leading to partial restoration of a certain amount of wild-type *IDS* RNA and protein which appeared to have, at least to some extent, an appropriate lysosomal localization; somewhat unexpectedly, this occurred irrespective of the cell type employed.

To further investigate the finding of a partial lysosomal localization, we extended the analysis to other prominent cellular bodies in the patients' fibroblasts using the confocal high-content screening system, which enables automatic high-content imaging and analysis of thousands of individual cells visualized by high resolution microscopy. The screening performed on lysosomes, the Golgi apparatus and early-endosomes confirmed adequate, although somewhat low, lysosomal localization (approximately 10%) of the proteins encoded by the mutant mRNAs (p.P4SfsX43 and p.R8X) edited-to-wild-type and clearly demonstrated that they did not accumulate in the Golgi nor in the early endosomes of the two patients' cell lines (Figure 7 C-D).

Overall, these results concurred with those we previously reported in the Hunter syndrome patient samples [Lualdi et al., 2010], but they still did not provide any obvious explanation for the deficiency of IDS enzymatic activity in our patient samples. The observed low residual enzymatic activity could in principle be explicable in terms of the preferential lysosomal degradation of the 'corrected' proteins as demonstrated in a recent study of RNA editing in mammalian brain cells [Behm and Öhman, 2016].

Formatted: Normal, Tab stops: Not at 0.29" + 0.39"

The only other report of an editing-like variant-correction mechanism occurring in the context of a human disease gene has been that tentatively noted in the context of *GNPTG* [Velho et al., 2016]. A homozygous c.[328G>T] transversion predicting a p.[Glu110Term] substitution was detected in siblings affected by another lysosomal disorder, mucolipidosis III gamma, but this variant was undetectable at the cDNA level (only the wild-type *GNPTG* sequence was evident). Although this additional example of 'enigmatic transcript correction' remains to be confirmed by in-depth analysis, we find it intriguing that it is also associated with a lysosomal disorder, particularly in the light of our above suggestion of the preferential lysosomal degradation of the 'corrected' proteins.

The possible instructive role of the nucleotide sequence environment around the mRNA editing sites remains to be clarified. Our findings demonstrate that the *ex vivo* repair mechanism occurring in all cell types tested is likely to be critically dependent upon the intrinsic characteristics of the primary sequence/secondary structure of the mRNA editing sites themselves. Some proteins involved in the response to DNA damage may function in RNA metabolism [Kai, 2016]; conversely, RNA processing proteins or protein complexes may be involved in the DNA damage response or may even exhibit DNA repair activity themselves [Jobert and Nilsen, 2014]. RNA processing enzymes are known to play a role in the maintenance of genomic integrity and in the integration and interconnection of DNA and RNA processing pathways in the eukaryotic nucleus. Currently available data point to an emerging role for specific Base Excision Repair (BER) proteins in RNA metabolism and RNA surveillance pathways [Jobert and Nilsen, 2014]. The unique ability of BER DNA glycosylases to recognize even subtle chemical modifications of nucleic acid bases may serve to distinguish normal and aberrant RNA molecules [Jobert and Nilsen, 2014]. We may therefore speculate that proteins involved in the signaling and repair of DNA damage might also function as part of a novel mRNA correction system. Such a system could in principle be recruited to the task of experimentally correcting inherited gene lesions at the RNA level, thereby opening up new therapeutic avenues for the future.

We conclude that despite the editing of the mutant *IDS* mRNA molecules to wild-type, and despite these mRNAs being translated, the fate of the protein products of translation of these mRNAs is qualitatively and quantitatively different from that expected of normal wild-type IDS. We rationalize the situation as follows: firstly, the *IDS* mRNA transcripts harbouring the nonsense/frameshift mutations have managed to elude nonsense-mediated mRNA decay. Secondly, a sizeable proportion of the mutant *IDS* mRNA transcripts have been edited to wild-type. Thirdly, despite the correction of the *IDS* mRNA sequence by editing so that they can be properly translated, the cell still appears to treat 'corrected' IDS protein translated from these mRNAs differently from IDS protein translated from wild-type *IDS* mRNAs in normal individuals. Thus, the success of the mRNA editing correction process notwithstanding, it may well be that these essentially wild-type *IDS* mRNAs have already been tagged by the cell as 'abnormal' and hence all subsequent translation products have, at an early stage, been earmarked for destruction despite the evident translatability of the edited *IDS* mRNAs [Inada, 2017].

Formatted: Font: Italic

Formatted: Font: Italic

Formatted: Font: Italic

Formatted: Font: Italic

Formatted: Font: Italic

Formatted: Font: Italic

Acknowledgments

Patient samples were obtained from the "Cell Line and DNA Biobank from patients affected by Genetic Diseases" (Istituto Giannina Gaslini), member of Telethon Network of Genetic Biobanks (project no. GTB12001). This work was partially supported by unrestricted grants from "Cinque per mille e Ricerca Corrente, "Ministero della Salute".

The authors acknowledge the technical assistance of, and valuable discussions with, Dr. Tiziana Bachetti from 'UOC Genetica Medica, Istituto Giannina Gaslini'.

The authors are unaware of any competing interests

REFERENCES

Aphasizhev R, Aphasizheva I. 2011. Uridine insertion/deletion editing in trypanosomes: a playground for RNA-guided information transfer. Wiley Interdiscip Rev RNA 2:669-85.

Aphasizheva I, Aphasizhev R. 2016. U-insertion/deletion mRNA-editing holoenzyme: definition in sight. Trends Parasitol 32:144-56.

Bass BL. 2002. RNA editing by adenosine deaminases that act on RNA. Annu Rev Biochem 71:817-46.

Bazak L, Haviv A, Barak M, Jacob-Hirsch J, Deng P, Zhang R, Isaacs FJ, Rechavi G, Li JB, Eisenberg E, Levanon EY. 2014. A-to-I RNA editing occurs at over a hundred million genomic sites, located in a majority of human genes. Genome Res 24:365-76.

Behm M and Öhman M. 2016. RNA editing: A contributor to neuronal dynamics in the mammalian brain. Trends Genet 32:165-75.

Blow MJ, Grocock RJ, van Dongen S, Enright AJ, Dicks E, Futreal PA, Wooster R, Stratton MR. 2006. RNA editing of human microRNAs. Genome Biol 7:R27.

Bruschi M, Musante L, Candiano G, Ghiggeri GM, Herbert B, Antonucci F, Righetti PG. 2003. Soft immobilized pH gradient gels in proteome analysis: a follow-up. Proteomics 3:821-5.

Bruschi M, Seppi C, Arena S, Musante L, Santucci L, Balduini C, Scaloni A, Lanciotti M, Righetti PG, Candiano G. 2005. Proteomic analysis of erythrocyte membranes by soft Immobiline gels combined with differential protein extraction. J Proteome Res 4:1304-9.

Chen C, Bundschuh R. 2012. Systematic investigation of insertional and deletional RNA-DNA differences in the human transcriptome. BMC Genomics 13:616.

Chen JY, Peng Z, Zhang R, Yang XZ, Tan BC, Fang H, Liu CJ, Shi M, Ye ZQ, Zhang YE, Deng M, Zhang X, et al. 2014. RNA editome in rhesus macaque shaped by purifying selection. PLoS Genet 10:10.

Chen L. 2013. Characterization and comparison of human nuclear and cytosolic editomes. Proc Natl Acad Sci USA 110:E2741-7.

Cruz-Reyes J, Zhelonkina A, Rusche L, Sollner-Webb B. 2001. Trypanosome RNA editing: simple guide RNA features enhance U deletion 100-fold. Mol Cell Biol 21:884-92.

Daniel C, Silberberg G, Behm M, Ohman M. 2014. Alu elements shape the primate transcriptome by cis-regulation of RNA editing. Genome Biol 15:R28.

den Dunnen JT, Dalgleish R, Maglott DR, Hart RK, Greenblatt MS, McGowan-Jordan J, Roux AF, Smith T, Antonarakis SE, Taschner PE. 2016. HGVS recommendations for the description of sequence variants: 2016 Update. Hum Mutat 37:564-9.

Estévez AM, Simpson L. 1999. Uridine insertion/deletion RNA editing in trypanosome mitochondria - a review. Gene 240:247-60.

Filocamo M, Mazzotti R, Corsolini F, Stroppiano M, Stroppiana G, Grossi S, Lualdi S, Tappino B, Lanza F, Galotto S and Biancheri R. 2014. Cell Line and DNA Biobank From Patients Affected by

Formatted: German (Germany)

Genetic Diseases. Open Journal of Bioresources 1:e2, DOI: http://dx.doi.org/10.5334/ojb.ab

Gott JM. 2011. RNA editing and human disorders. In: eLS. John Wiley & Sons, Ltd: Chichester doi:10.1002/9780470015902.a0005494.

Griffiths-Jones S, Saini HK, van Dongen S, Enright AJ. 2008. miRBase: tools for microRNA genomics. Nucleic Acids Res 36:D154-8.

Grohmann M, Hammer P, Walther M, Paulmann N, Büttner A, Eisenmenger W, Baghai TC, Schüle C, Rupprecht R, Bader M, Bondy B, Zill P, et al. 2010. Alternative splicing and extensive RNA editing of human *TPH2* transcripts. PLoS ONE 5:e8956.

Harjanto D, Papamarkou T, Oates CJ, Rayon-Estrada V, Papavasiliou FN, Papavasiliou A. 2016. RNA editing generates cellular subsets with diverse sequence within populations. Nat Commun 7:12145.

Hwang T, Park CK, Leung AK, Gao Y, Hyde TM, Kleinman JE, Rajpurohit A, Tao R, Shin JH, Weinberger DR. 2016. Dynamic regulation of RNA editing in human brain development and diseases. Nat Neurosci 19:1093-9.

Inada T. 2017. The ribosome as a platform for mRNA and nascent polypeptide quality control. Trends Biochem Sci. 42:5-15.

Jobert L and Nilsen H. 2014. Regulatory mechanisms of RNA function: emerging roles of DNA repair enzymes. Cell Mol Life Sci 71:2451-65.

Kai M. 2016. Roles of RNA-binding proteins in DNA damage response. Int J Mol Sci 17:310.

Kim MS, Hur B, Kim S. 2016. RDDpred: a condition-specific RNA-editing prediction model from RNA-seq data. BMC Genomics 11 Suppl 17:1-5.

Kleinman CL, Adoue V, Majewsk J. 2012. RNA editing of protein sequences: a rare event in human

Formatted: Font: Italic

Formatted: Spanish (Spain)

transcriptomes. RNA 18:1586-96.

Li M, Wang IX, Li Y, Bruzel A, Richards AL, Toung JM, Cheung VG. 2011. Widespread RNA and DNA sequence differences in the human transcriptome. Science 333:53-8.

Li S, Maso, CE. 2014. The pivotal regulatory landscape of RNA modifications. Annu Rev Genomics Hum Genet 15:127-50.

Lualdi S, Tappino B, Di Duca M, Dardis A, Anderson CJ, Biassoni R, Thompson PW, Corsolini F, Di Rocco M, Bembi B, Regis S, Cooper DN, et al. 2010. Enigmatic *in vivo* iduronate-2-sulfatase (IDS) mutant transcript correction to wild-type in Hunter syndrome. Hum Mutat 31:E1261-85.

Maas S, Kawahara Y, Tamburro KM, Nishikura K. 2006. A-to-I RNA editing and human disease. RNA Biol 3:1-9.

Mrowka C, Schedl A. 2000. Wilms' tumor suppressor gene *WT1*: from structure to renal pathophysiologic features. J Am Soc Nephrol 11 Suppl 16:S106-15.

Navaratnam N, Sarwar R. 2006. An overview of cytidine deaminases. Int J Hematol 83:195-200.

Niavarani A, Currie E, Reyal Y, Anjos-Afonso F, Horswell S, Griessinger E, Luis Sardina J, Bonnet D. 2015. APOBEC3A is implicated in a novel class of G-to-A mRNA editing in WT1 transcripts. PLoS One 10:e0120089.

Osato D, Rogers K, Guo Q, Li F, Richmond G, Klug F, Simpson L. 2009. Uridine insertion/deletion RNA editing in trypanosomatid mitochondria: In search of the editosome. RNA 15:1338-44.

Peng Z, Cheng Y, Tan BC, Kang L, Tian Z, Zhu Y, , Zhang W, Liang Y, Hu X, Tan X, Guo J, Dong Z, et al. 2012. Comprehensive analysis of RNA-Seq data reveals extensive RNA editing in a human transcriptome. Nat Biotechnol 30:253-60.

Formatted: Font: Italic

Formatted: Spanish (Spain)

Formatted: Font: Italic

Formatted: Spanish (Spain)

Piskol R, Ramaswami G, Li JB. 2013. Reliable identification of genomic variants from RNA-seq data. Am J Hum Genet 93:641-51.

Ramaswami G, Li JB. 2016. Identification of human RNA editing sites: A historical perspective. Methods 107:42-47.

Reautschnig P, Vogel P, Stafforst T. 2016. The notorious R.N.A. in the spotlight - drug or target for the treatment of disease. RNA Biol14:1-18.

Rieder LE, Staber CJ, Hoopengardner B, Reenan RA. 2013. Tertiary structural elements determine the extent and specificity of messenger RNA editing. Nat Commun 4:2232.

Savva YA, Reenan RA. 2014. Identification of evolutionarily meaningful information within the mammalian RNA editing landscape. Genome Biol 15:103.

Sharma PM, Bowman M, Madden SL, Rauscher FJ, Sukumar S. 1994. RNA editing in the Wilms' tumor susceptibility gene, *WT1*. Genes Dev 8:720-31.

Slotkin W, Nishikura K. 2013. Adenosine-to-inosine RNA editing and human disease. Genome Med 5:105.

Stenson PD, Mort M, Ball EV, Evans K, Hayden M, Heywood S, Hussain M, Phillips AD, Cooper DN. 2017. The Human Gene Mutation Database: towards a comprehensive repository of inherited mutation data for medical research, genetic diagnosis and next-generation sequencing studies. Hum Genet. 2017 Mar 27. doi: 10.1007/s00439-017-1779-6. [Epub ahead of print]

Velho RV, Ludwig NF, Alegra T, Sperb-Ludwig F, Guarany NR, Matte U, Schwartz IV. 2016. Enigmatic *in vivo* GlcNAc-1-phosphotransferase (GNPTG) transcript correction to wild type in two mucolipidosis III gamma siblings homozygous for nonsense mutations. J Hum Genet 61:555-60.

Formatted: German (Germany)

Formatted: Font: Italic

Formatted: Font: (Default) Times New Roman, 12 pt, Not Bold, Complex Script Font: Times New Roman, 12 pt, English (United States)

Formatted: Font: (Default) Times New Roman, 12 pt, Complex Script Font: Times New Roman, 12 pt, English (United States)

Formatted: Font: (Default) Times New Roman, 12 pt, Not Bold, Complex Script Font: Times New Roman, 12 pt

Formatted: Font: (Default) Times New Roman, 12 pt, Complex Script Font: Times New Roman, 12 pt, English (United States)

Formatted: English (United States)

Formatted: Font: Italic

Villegas J, Müller I, Arredondo J, Pinto R, Burzio LO. 2002. A putative RNA editing from U to C in a mouse mitochondrial transcript. Nucleic Acids Res 30:1895-1901.

Wedekind JE, Dance GS, Sowden MP, Smith HC. 2003. Messenger RNA editing in mammals: new members of the APOBEC family seeking roles in the family business. Trends Genet 19:207-16.

Xu G, Zhang J. 2014. Human coding RNA editing is generally non adaptive. Proc Natl Acad Sci USA 111:3769-74.

Zhu H, Urban DJ, Blashka J, McPheeters MT, Kroeze WK, Mieczkowski P, Overholser JC, Jurjus GJ, Dieter L, Mahajan GJ, Rajkowska G, Wang Z, et al. 2012. Quantitative analysis of focused A-to-I RNA editing sites by ultra-high-throughput sequencing in psychiatric disorders. PLoS ONE 7:e43227.

Zougman A, Ziółkowski P, Mann M, Wiśniewski JR. 2008. Evidence for insertional RNA editing in humans. Curr Biol 18:1760-5.

FIGURE LEGENDS

Figure 1. Representative western blots on lymphoblasts from Hunter syndrome patients (p.P4Sfs and p.R8X) and controls (WT). IDS protein detection under SDS-reducing (A) and native (B) conditions. The western blots, obtained using an anti-IDS antibody, exhibited a major band in samples from both controls and patients, which exhibited identical electrophoretic migration properties to that of the IDS protein reference noted in SDS-PAGE. (C) Bar plot of IDS protein expression levels. The values from the densitometric analysis are expressed as optical density (OD) and revealed a decrease in IDS protein expression between the control and the patient harboring p.P4Sfs. The values quoted represent the means of triplicate measurements. Error bars represent standard deviations. Anti-tubulin was used as a housekeeping control.

Figure 2. Representative two-dimensional gel electrophoresis of immunoprecipitation from Hunter syndrome patients (p.P4Sfs and p.R8X) and controls (WT). Representative silver staining of the anti-IDS immunoprecipitation revealing three spots (circled) corresponding to the molecular weight of IDS protein. These three spots exhibited optical density changes comparable to those noted in the western blot experiments; however, no difference in either molecular weight or isoelectric point was evident between the control and the patient harboring p.R8X. It should be noted that the p.P4Sfs sample is not visible, probably due to its low expression (also observed by western blotting).

Figure 3. Capillary electrophoretic patterns obtained from transfection of expression vectors. The expression vectors contained (P4Sfs)-, (R8X)- and (W12X)-IDS cDNA (A, B and C, respectively) into patient fibroblasts (a,b), normal control fibroblasts (c) and the HEK293 (d) and HeLa (e) cell lines. Detection primers were designed so that their 3' ends corresponded to the nucleotides (nt) closest to the uracils (thymines) to be corrected; that is at the –1 position (forward primer) for the single base deletion, and at the +2 position (reverse primer) for the base conversion [Lualdi et al, 2010]. Either ddCTP (base deletion, black peak) or ddGTP (base conversion, blue peak) would be incorporated when the mutant *IDS* mRNAs were corrected, whereas ddTTP and ddATP

would be added if the mutant *IDS* mRNAs remained uncorrected. In the case of base deletion, the black peak would denote the normal nt C, whilst the red peak would denote the inserted (unedited) nt T (Figure 3A). In the case of base conversion (detection primer being the reverse), the blue peak would denote the normal (edited) nt G (C), whereas the green peak would denote the mutated (unedited) nt A (T) (Figure 3B). In addition, in the case of G>A conversion (W12X), using a reverse primer, only ddATP (red peak) was added as the mutant *IDS* mRNA remained uncorrected (Figure 3C).

Figure 4. **IDS-GFP protein expression in HEK293 cells and co-localization with lysosomes visualised by confocal microscopy**. HEK293 cells, transiently transfected with wild-type (A) or mutated (B and C) IDS-GFP (*green fluorescence*), were loaded with Hoechst 33342 dye (*blue fluorescence*) to stain the nucleus and incubated with LysoTracker (*red fluorescence*) to stain the lysosomes. The cells were immediately analysed by confocal microscopy as described in the Materials and Methods. The cells transfected with either IDS(WT)-GFP or with IDS(R8X)-GFP expressed a high level of chimeric protein, whereas the cells transfected with IDS(W12X)-GFP displayed almost no GFP signal (second column, A, B and C, respectively). In the third column, a cytoplasmic distribution which left the nucleus free of staining was demonstrated by LysoTracker staining (*red*). The fourth column (Merge) displays the distribution of IDS-GFP (*green*) in cells labeled with LysoTracker (*red*) showing localization within lysosomes (*yellow-orange*): confocal images showed significant overlap (*yellow-orange*) for IDS(WT)-GFP(A) and IDS(R8X)-GFP (B) but no overlap for IDS(W12X) (C). Colocalization with LysoTracker is particularly clear in cells expressing low levels of IDS-GFP (arrowheads), as in the inset with higher magnification in (B). In the fifth column, differential interference contrast (DIC) images show the number and morphology of the cells expressing IDS(WT), IDS(R8X) and IDS(W12X).

Figure 5: Flow cytometric analysis of GFP expression. (A) Dot plot of HEK293 Side Scatter (SSC) peak width against SSC peak height set to gate on single cells and to exclude aggregates. (B) Single

cell population analyzed for GFP level of expression. Considered a slight auto-fluorescence of cultured cells in the first emission channel (FL1), the negative is set on CTRL (not expressing GFP) auto-fluorescence. HEK 293 cells harbouring IDS(R8X) display a level of GFP expression (green spots on the right) that almost overlaps with the IDS(WT) sample (57% and 54%, respectively). In HEK293 cells harbouring IDS(W12X), the GFP expression is strongly reduced (16%) whilst the mean fluorescence intensity (MFI) is distinctly lower.

Figure 6. **Image flow cytometric co-localization analysis**. (A) The histogram values derive from a feature (Gradient RMS) that measures and plots the sharpness of an image, thereby determining its overall focus quality. In-focus cells are gated and then (B) analyzed on the basis of two physical qualities: Aspect ratio and Area of the brightfield. Aspect ratio was calculated based on brightfield as the ratio of cellular minor axis (width) to major axis (height). Round cells have an aspect ratio close to 0.8-1, whereas the elongated cells or clumps have a lower aspect ratio. We gated on round single cells of the correct dimension (Area) for further analysis (gate "HEK293"). (C) HEK293 cells were visualized on a dot plot based on their expression of GFP and LysoTracker. (D) To improve the visual observations and identify the best images, a statistical analysis of double positive events similarity in the staining patterns was performed. The *bright detail similarity* values were calculated using an algorithm for co-localization based upon the relative intensities in individual pixels of a cell. Since each pixel has a spatial registry, pixel intensity values in different channels can be compared. IDS(WT) (E) and IDS(R8X) (F) co-localization are shown (*yellow-orange in Ch02/Ch05*).

Figure 7. Representative images obtained from confocal high-content analysis of cells expressing wild-type or mutated IDS. IDS protein (red) immunolocalization in fibroblasts from control individuals (wt-IDS) and from Hunter syndrome patients (P4Sfs-, R8X- and W12X-IDS). (A) Wild-type IDS protein is expressed in fibroblasts from a healthy individual and co-localizes with LAMP-1, a lysosome marker (green). (B) In fibroblasts from Hunter syndrome patients bearing the p.W12X mutation, mutant IDS is expressed at a very low level, and does not co-localize with LAMP-

1. Fibroblasts from Hunter syndrome patients bearing either the p.R8X (C) or the p.P4Sfs (D) mutations exhibit very high levels of expression of the edited mutant protein, which partially co-localizes with LAMP-1 (lysosome). However, no co-localization is evident with the Golgi marker, giantin (yellow) or the early endosome marker, EEA1 (yellow). Bar = 10 µm.