Supplemental Methods

AAVrh.10hCLN2 administration

All children receiving the vector were prepared for anesthesia and surgery in a standard fashion. Design of the sites for vector administration specifically for each child was carried out within 24 hr prior to vector administration using the Brainlab system for image-guided surgery (Brainlab) based on a pre-operative MRI scan with head sentinels. The burr holes were made at the pre-determined marked locations, the dura opened, and 150 μm diameter flexible glass catheters (Polymicro Technologies) used to administer the vector (*52, 53*). A 20-gauge spinal needle was placed on the surface of the brain orthogonal to the skull to act as a guide for catheter insertion 2 cm into the pre-determined locations. Intravenous mannitol (typically 1.0 g/kg) was given as needed throughout the period of vector administration to minimize brain edema. In all treated children, the total vector volume of 1.8 ml was equally divided among 12 cortical locations delivered through 6 burr holes (2 locations at 2 depths through each burr hole), 3 burr holes per hemisphere. While the exact locations of the administration of the vector were subject specific, they were generally in the same regions of the brain, with the goal of providing safe, widespread distribution as has been previously described (*53*). Briefly, six trajectories were planned for each subject, with 3 bilateral paired trajectories targeting the subcortical white matter and entering through the middle of the superior frontal gyrus, immediately anterior to the precentral gyrus and the posterior superior parietal lobule. Deep targets for each hemisphere relative to the mid-commissural point (midpoint of the intercommissural line between the anterior and posterior commissures) was as follows:

Anterior frontal: X=25 mm lateral, Y=30 mm anterior, Z=25 mm superior

Posterior frontal: X= 20 mm lateral, Y=2 mm anterior, Z= 30 mm superior

Parietal: X=20 mm lateral, Y=35 mm posterior, Z=30 mm superior

Entry points were planned at the closest gyrus which was perpendicular to the planned target with the goal of the deep target being roughly 25 mm below the cortical surface. To create perpendicular tracts that would minimize the risk of an angular trajectory skewing into a sulcus, the deep target was adjusted up to 5 mm in any direction, also ensuring that the deep target was within the white matter below the bottoms of the adjacent sulci to facilitate wider spread of the vector solution, as described previously (53). Following administration of vector to the deep target at each of the six bilateral locations, catheters were withdrawn roughly 10 mm where a second infusion was completed at each site. An example of surgical planning for subject V1 is provided (fig. S2). The vector was administered at a rate of 2.0 μl/min to each of the 6 sites (the deeper of the 2 sites through each burr hole) simultaneously by a microperfusion pump (Harvard Instrument PHD 2000 Infuse/Withdraw Multichannel Syringe Pump, Harvard Apparatus, Holliston, MA) driving 6 Hamilton syringes (Hamilton Syringe, Reno, NV). After the specified dose was administered, the catheters were left in place for 5 min to assure tissue penetration. The catheters were then withdrawn approximately half-way from the bottom of the catheter tract to the brain surface, and the remaining 50% of the dose was administered, in parallel, to each of the 6 additional sites as described above. For both cohorts 1 and 4, the average time for total vector infusion averaged 151.1±1.0 min, surgery duration was 358±43 min and time under anesthesia was 481±55 min (see tables S2 and S3 for details). Following vector administration, the incision was closed with standard techniques. A post-operative CNS MRI was performed within the first 48 hr following the surgical procedure to assess for bleeding or other possible peri-operative adverse events.

The first 6 children in cohort 1 (V1 to V6) received a total dose of 9.0×10^{11} genome copies (gc) delivered in equal doses at each of 12 sites (7.5×10^{10} gc per site). In some children, foci of T2 hyperintensity localized to the sites of administration were observed in the day 1 post-

surgical T2 FLAIR, diffusion weighted imaging (DWI), and apparent diffusion coefficient (ADC) MRI. These abnormalities persisted in subsequent CNS MRIs (6 to 12 months post-administration) in most cases, while in others it resolved (see fig. 1 and fig. S3 for examples). The volume of hyperintensity was estimated on the T2-FLAIR MRI images by defining a region of interest (ROI) that outlined each area of increased signal intensity on all slices. The total number of voxels in the ROI was determined and multiplied by the volume of each voxel to produce a total volume (table S5). Although there were no clinical correlates attributable to the persistent localized foci of T2 hyperintensity, in agreement with the FDA, IRB and DSMB, the dose for subsequent children (V7, V8 in cohort 1 and S3, S4, S5 in cohort 4) was reduced to a total dose of 2.85×10^{11} gc (2.4×10^{10} gc per site). Comparisons within each time point, pre and post vector administration, between dose groups were made using the Fisher exact test in a 2 x 2 table (High-low dose *vs.* number of positive and negative abnormalities) for each parameter (T2, DWI and ADC).

Post-vector administration assessment

Each child was monitored post-operatively in the recovery room and pediatric intensive care unit, and once stable, transferred to an inpatient bed. The children were discharged from the hospital an average 3.0 ± 1.0 days post-surgery. All families were asked to remain in the proximity of the hospital until the day 14 evaluation.

Children in cohorts 1 and 4 were assessed at Weill Cornell at days 7 and 14, and at months 1, 6, 12, and 18 following treatment for safety parameters. At month 2 and 3, they were additionally assessed for adverse effects at the child's personal physician's office (see table S11 for timeline of safety assessments). For one subject (V6), who was unable to return to Weill Cornell for follow up visits, the study team went to the subject's home location to carry out some of the follow-up visits. All clinical efficacy evaluations for cohort 1 using the clinical rating scale

were videotaped for blinded assessment by 3 pediatric neurologists (*4*) (table S7). The family for subject V8 dropped out of the 18-month follow-up part of the study, citing difficulty in traveling with the child. At month 22 after therapy, we sent a team to assess the child. Because none of the assessments were within the mandated 18 month ± 30 day study period (table S11), the data was not used for efficacy analysis.

Safety parameters

The safety parameters were assessed over the course of the 18 months for both cohorts 1 and 4 (table S11). Adverse event information was captured and the clinical monitor determined the attribution of adverse events to the study drug. Based on prior experience indicating possible localized inflammation and/or edema at the sites of administration when the vector concentration at the tip of the catheters are the highest, the CNS MRI pre- and post-administration (days 1, months 6, 12 and 18) were assessed for the presence of T2 FLAIR and diffusion abnormalities at the estimated sites of administration (fig. 1).

Anti-vector immunity

Serum AAVrh.10 neutralizing antibody titers from cohorts 1 and 4 over time were quantified by an in vitro assay with HEK293-ORF6 cells. An AAVrh.10 vector expressing a luciferase reporter transgene (AAVrh.10Luc) was incubated with 2-fold serial dilutions of sera at 37°C for 45 min and then used to infect cells at a multiplicity of infection of 3000 genome copies/cell. At 48 hr post-infection, luciferase activity was assessed with cell lysate using the Luciferase Assay System (Promega, Madison, WI). The neutralizing antibody titer was expressed as the reciprocal of serum dilution at which 50% inhibition of AAVrh.10Luc was observed (*58*). Similarly, CSF AAVrh.10 neutralizing antibody titers from cohort 1 over time (on pre and one post administration timepoint) were quantified by an in vitro assay with HEK293-ORF6 cells as described above for the serum samples.

For anti-capsid and anti-transgene cellular immunity, blood samples were collected from cohort 1 and cohort 4 at timepoints specified in the timeline (table S11), fractionated and sent to the Immunology Core, Gene Therapy Program at the University of Pennsylvania. Isolated peripheral blood mononuclear cells (PBMC) were assayed for T-cell responses to the AAVrh.10 capsid and CLN2 transgene by INF-γ ELISpot with 3 pools of AAVrh.10 capsid peptides and 2 pools of transgene peptides, each synthesized as libraries of 15-mers with a 10 amino acid overlap (Mimotopes). As a control, the potential for toxicity of these peptides was evaluated for the inhibition of a stimulated response in a standard blood mononuclear cell preparation to a positive control peptide library, a panel of MHC class 1 restricted viral peptides from cytomegalovirus, Epstein-Barr virus and influenza virus (*59*). Stimulation with phytohemagglutinin (PHA) provided the positive assay control; the negative control was growth media. The number of spot-forming units (sfu) per 10^6 PMBC was counted. Data accepted as valid included only samples that had positive PHA response and low sfu for the media stimulated control.

Relative quantity of CSF TPP1

Human TPP1 protein expression was assessed in cerebral spinal fluid (CSF) collected at pre- and one post-administration follow-up visit, (the Weill Cornell IRB restricted post-therapy CSF assessment to 1 time-point). CSF from 3 healthy children were mixed in equal volume and 1-10 µl were analyzed to serve as positive control. CSF from the study children both pre- and post-vector administration (10 µl) was analyzed in a 4 to 12 % SDS-polyacrylamide gel and transferred onto a polyvinylidene difluoride membrane. The membrane was treated with rabbit anti-human TPP1 antibody, 1:1000 diluted in 5 % dry milk in PBS, (Abcam) for 1 hr, 23°C and then washed 4 times with PBS plus 0.05 % Tween-20 (PBS-Tween). The membrane was then incubated with 1:5,000 diluted horseradish peroxidase (HRP)-conjugated goat anti-rabbit IgG

(Abcam) for 1 hr, 23°C, washed 5 times with PBS-Tween and developed with Enhanced

Chemiluminescence (ECL) Plus reagent (Thermo Scientific). The amount of TPP1 was

quantified using Image J software and expressed as integrated band density in arbitrary units.

Fold-increase (relative quantity) of TPP1 l following vector administration was compared to the

pre-administration amount of TPP-I for the same child. The % normal amount was determined

by using the TPP-I band density in the linear range of normal levels normalized to 10 μl.

CNS MRI % grey matter

As controls for the MRI % grey matter assessment of the treated children, 62 MRI

datasets were acquired from 47 untreated CLN2 patients. The control data included 12 untreated

controls (cohort 2), the pre-therapy time points for the n=8 children in cohort 1 and n=5 in cohort

4 (the "screening" and "pre-therapy" assessments; table S11) and n=24 children in the screening

study that did not participate in the study. For the treated children in cohort 1, there were 3 post-

therapy MRI evaluations, at 6, 12 and 18 months.

All imaging data were acquired on a 3.0 Tesla GE MRI scanner with an 8-channel head

coil. A sagittal BRAVO sequence was used with isotropic (1.0 x 1.0 x 1.0 mm) resolution as

previously described (*36*). Percent gray matter (%GM; % of total brain volume) was calculated

using the FAST segmentation program within the FSL Software Library (FMRIB, Oxford UK)

(*60*).

The skull was digitally removed prior to segmentation using the FSL brain extraction tool

(*61*). The %GM was determined by multiplying the mean value of the tissue probability by the

tissue volume and dividing by the total of gray matter + white matter (WM) + CSF (*62*). A

sigmoidal function was tested for the imaging variables of the untreated children as defined by:

$$\%GM = \frac{A_1}{1+e^{A_2(Age-A_3)}} + A_4$$

where A_1 is the amplitude of the sigmoidal curve, A_2 determines the sharpness of the rate of decline, A_3 is the time shift and A_4 is the decay asymptote. Since the age of onset of CLN2 is variable, and thus age is not an independent variable, the data was fitted using a total least squares method. Unlike conventional least squares regression models, total least squares regression is not scale invariant, and so requires a scaling rule to be specified. Based upon inspection of the %GM dataset, we chose $\%GM_{scaled} = 27.05 * \%GM - 4.66$, such that the numerical range of $\%GM_{scaled}$ approximated the range of subject ages in years in order to make the dynamic range of the x and y axis roughly equivalent. After fitting the natural history cohort using this method, the scaling factor was removed.

Fitting was performed using a bootstrap technique. For each bootstrap, the 62 points from the CLN2 natural history dataset were resampled with replacement, i.e. duplicates were allowed, using MATLAB 2019a (Mathworks). This process was repeated 1000 times. Next, the results from each run of the bootstrap were fitted with the sigmoid of Equation 1, consisting of 100 points across an age range of 2 to 12 yr. After completion of 1000 runs, the $\%GM_{scaled}$ mean value and 95% confidence intervals were calculated for each of the 100 points. The resulting functions were closely approximated by sigmoids as expected.

In order to eliminate the time variable from the analysis, the rate of change of %GM was plotted *vs* %GM, allowing for direct comparison of subjects with different ages of disease onset. If the %GM varies with age according to the sigmoid of equation 1, then the %GM/yr *vs* %GM is a parabolic function. Since the parabola was calculated by first taking the time derivative of the sigmoid, and therefore each data point represents the difference of the neighboring data-points of the %GM sigmoid, the error in the parabolic function at each value of %GM was estimated simply from the quadrature sum of errors in the sigmoid as:

$$\delta\%GM/yr(t) = sqrt\ (\delta\%GM(t)^2 + \delta\%GM(t\text{-}1)^2\ )$$

where δ%GM(t) is the difference between the sigmoidal function and its 95% confidence interval at time t.

Assessment of the effect of therapy *vs* the untreated controls was determined by comparing the difference between the 95% confidence intervals of the slopes of %GM decline in treated children in cohort 1 and the 95% confidence interval of the sigmoidal fit to CLN2 MRI natural history data at a given disease severity defined by the %GM. This method accounts for both ceiling and floor effects as estimated by the upper and lower asymptotes of the 95% confidence intervals of the sigmoids. Mean values of treated children above the 95% confidence interval was considered an improvement compared to the untreated controls. This data was further used to compare the MRI % grey matter change/yr with the range of change/yr for the untreated children with matching % grey matter.

Vision-related parameters

Given the severity of motor and cognitive abnormalities associated with CLN2 disease, all children were examined while under sedation. The baseline ophthalmic evaluation included complete anterior segment and dilated exam, fundus photography (RetCam, Clarity Medical Systems Inc), spectral domain optical coherence tomography (OCT, Heidelberg Engineering), fluorescein angiography (FA, Heidelberg Engineering and RetCam) and indocyanine green angiography (ICGA, Heidelberg Engineering). The ocular exam, OCT, FA, and ICGA were used to establish the extent of retinal degeneration in each patient based on the Weill Cornell Batten Scale as previously described (*39*). In addition to the baseline exam, 5 children underwent ophthalmic evaluation, with anterior segment and dilated exam and fundus photography, following gene therapy administration. In one of these children, OCT evaluation was also performed after therapy. In those children with OCTs, for each eye on each examination date, central subfield thickness (CST) was calculated by Heidelberg software as previously described

(*38*). In the 5 children with follow-up evaluation, the clinical exam and fundus photography were used as a qualitative assessment of retinal disease progression. Quantitative assessment of the progression of CLN2-related retinal degeneration in the one child with follow-up OCT was determined and compared to the natural history graph derived using CST as previous described (*38*).

Clinical neurologic rating scale

There are 2 clinical neurologic rating scales for CLN2 disease: the original Hamburg scale described by Steinfeld et al (*63*); and the Weill Cornell scale described by Worgall et al (*4*). Details of the 2 scales are described in table S1. The 2 scales are identical in 2 parameters: the "Motor" and "Language" parameters in the Hamburg scale are the same as the "Gait" and "Language" parameters in the Weill Cornell Scale. Each of these scales had additional parameters which were not used in assessment of the efficacy of the therapy because of variability, dependency on care giver parameters, or irrelevance to CNS disease. The "Motor" and "Language" (Hamburg scale) and the identical "Gait" and "Language" (Weill Cornell scale) were used to quantify the rate of clinical decline of the treated and untreated children. Since the "Gait" (Weill Cornell) and "Motor" (Hamburg) parameters are identical (table S1), to avoid confusion, we used the term "Motor" instead of "Gait" used in the Weill Cornell scale. As a comprehensive clinical neurologic assessment, the "motor" and "language" scores (each scale 0-3) were summed to generate a "CLN2 disease neurologic rating scale" (*64*). These were the identical parameters used in the multi-institutional collaborative CLN2 disease "natural history" publication (*5*), and the parameters for the FDA approval of cerliponase alfa (*31*).

For cohorts 1 and 2, the clinical assessment of motor + language was performed prospectively using defined standard operating procedures (SOPs) based on 3 to 4 observers, with specific rules on how the data was evaluated. The primary, on-site assessor was a pediatric

neurologist who had been trained on implementing the scale. The assessment of each child was videotaped by a trained technician following a SOP for recording the assessment and editing for review by 2 to 3 other pediatric neurologists who were trained on implementing the scale. The neurologists that assessed the video recording were blinded to the subjects' treatment status. In the event of discrepancy of more than 1 point between the 2 blinded scorers, a 3rd pediatric neurologist, also blinded, scored the video in order to act as a tie-breaker. The final score was an average of the assessment of the 3 to 4 reviewers (primary + 2 to 3 additional reviewers) to minimize bias and subjective interpretation. The variance among the observers was not significant (table S7). As a further validation of the methods used to assess the robustness of the quantitative neurologic assessment, reproducibility of the motor and language scale was validated by comparing repeat assessments of the severity of CLN2 disease in the same child carried out within <1.5 months, a time when deterioration would not be detectable. The mean assessment of the 3 to 4 observers was identical over this short time interval (table S5).

Individuals other than the principal and co-principal investigators collected, tabulated and verified the clinical parameters and adverse effects. To quantify the annual rate of decline of motor and language, linear regression of the consensus motor + language score over time was taken for each subject in the treatment (cohort 1) and untreated cohort (cohort 2). As the age of each subject was in days, the slope obtained was multiplied by 365 days to provide an annual rate of decline for each subject and the individual rates were then averaged to provide the annualized rate of decline ± standard deviation for the cohort. The individual data points for all subjects in cohorts 1 and 2 are in Table 2 and table S11.

Additional Safety Data

In addition to the treated cohort (cohort 1), there were 5 children (cohort 4) assessed in the screening protocol that did not fulfill the inclusion/exclusion criteria for the treated *vs*

untreated trial because of disease severity and/or genotype. At the request of the families, these

children were enrolled in a secondary safety study, under a different protocol where they

received the same therapeutic intervention but for the purpose only of adding to the safety profile

of AAVrh.10hCLN2 (NCT01161576, see table S10 for Inclusion/Exclusion criteria, Table 1 for

demographics and table S11 for timeline of assessments).

Secondary Parameters

The parents of all treated and untreated children were asked to complete the CHQ or

ITQoL (depending on age) quality of life questionnaires and the children were also assessed with

the Mullen scale (*65-68*). The Infant Toddler Quality of Life Questionnaire Parent Form (ITQoL-

PF97) was used to assess parents of 2 months to 5 year-old subjects while the Child Health

Questionnaire Parent Form (CHQ-PF50) was used to assess parents of 5 to 18 year-old subjects

(table S12). These quality of life questionnaires were completed by at least one parent/legal

guardian at the times of assessment. The survey was administered independently to each parent

to minimize observer bias if both parents were present. The Mullen pediatric developmental

psychological rating scale, was administered by either a neuropsychologist or trained study

coordinator. This scale assesses gross motor, cognitive, receptive and expressive language,

adaptive behavior and fine motor skills.
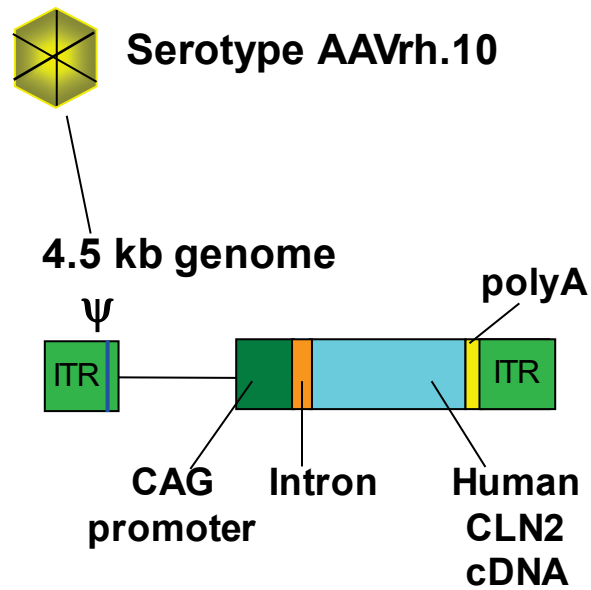
Figure S1. AAVrh.10hCLN2 vector. The vector is comprised of an AAVrh.10 capsid encompassing a genome composed of 5' and 3' AAV2 inverted terminal repeats surrounding an expression cassette including: the enhancer from human cytomegalovirus, promoter, splice donor and left hand intron sequence from chicken β-actin /right hand intron sequence and splice acceptor from rabbit β-globin, the normal human CLN2 cDNA, and rabbit β-globin polyA.
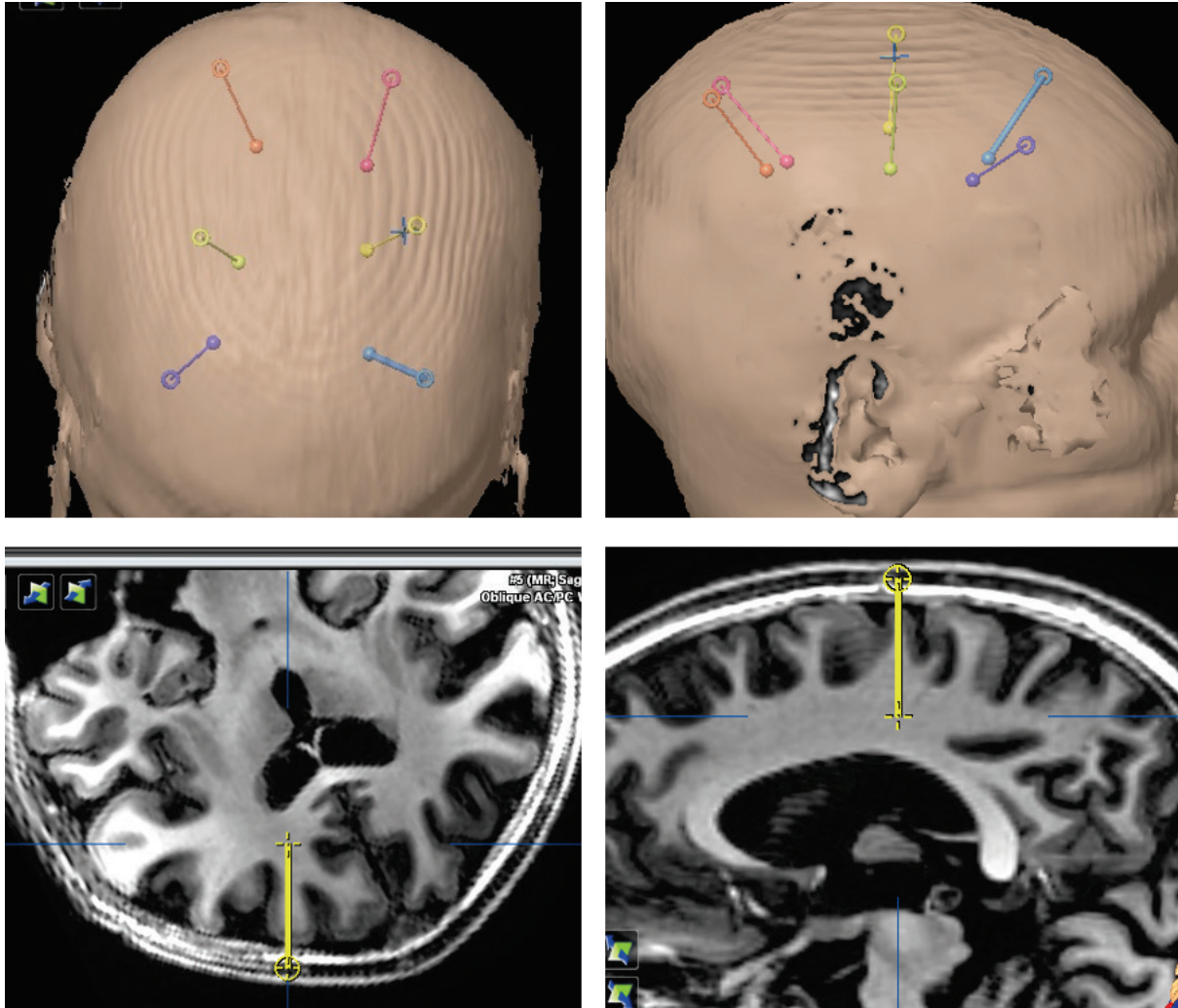
Figure S2. **Trajectory planning for gene therapy infusions.** Example of trajectory planning for a study subject showing six planned trajectories (bilateral anterior frontal, posterior frontal and parietal) overlaid on the reconstructed subject head in a frontal (upper left) and sagittal (upper right) views. An example of a single trajectory (left posterior frontal) from surface to the deep target shown through the long axis of the trajectory in oblique coronal (lower left) and sagittal (lower right) views.

T2 FLAIR 6 month   **T2 FLAIR 6 month**   **T2 FLAIR 6 month**

A. V2

B. V3

C. V7

D. V4

E. V5

F. V6
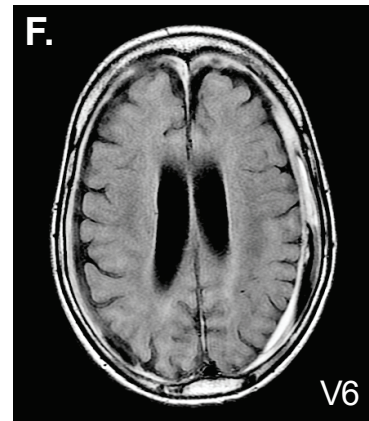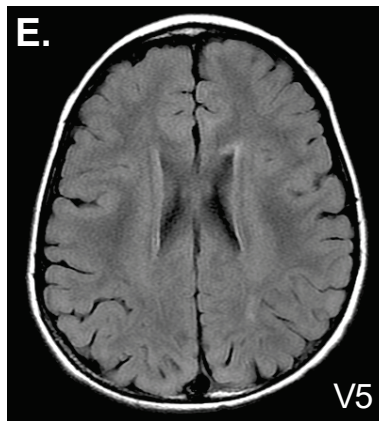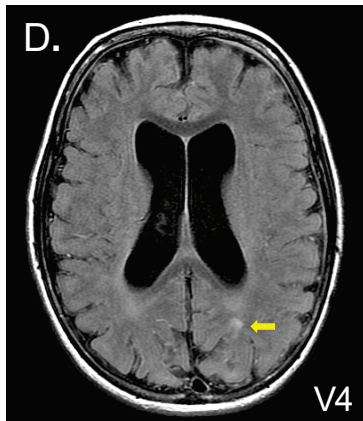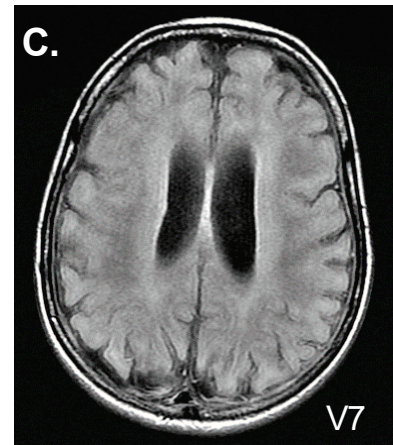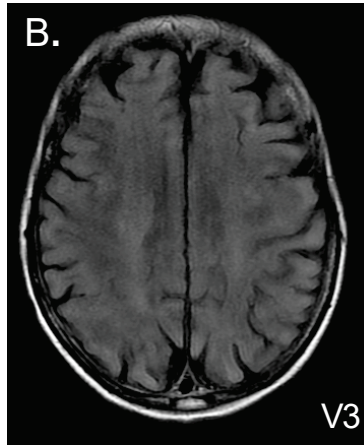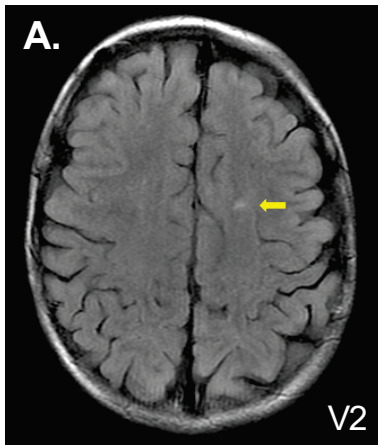
T2 FLAIR 12 month   **T2 FLAIR 12 month**   **T2 FLAIR 12 month**

Figure S3. **Axial T2 FLAIR (T2 FLAIR) assessment of participants post-therapy.**
Additional images of post-treatment examples of Axial T2 FLAIR (T2 FLAIR) MRI assessment
of subjects where the T2 hyperintensities observed were minimal or absent. **A.** Participant V2, 6
month post-administration, minimal T2 hyperintensity observed; **B.** V3, 6 months, no T2
hyperintensity observed; **C.** V7, 6 months, no T2 hyperintensity observed; **D.** V4, 12 months,
minimal T2 hyperintensity observed. **E.** V5, 12 months, no T2 hyperintensity observed. **F.** V6,
12 months, no T2 hyperintensity observed. Yellow arrows identify any abnormalities at the site
of vector administration.
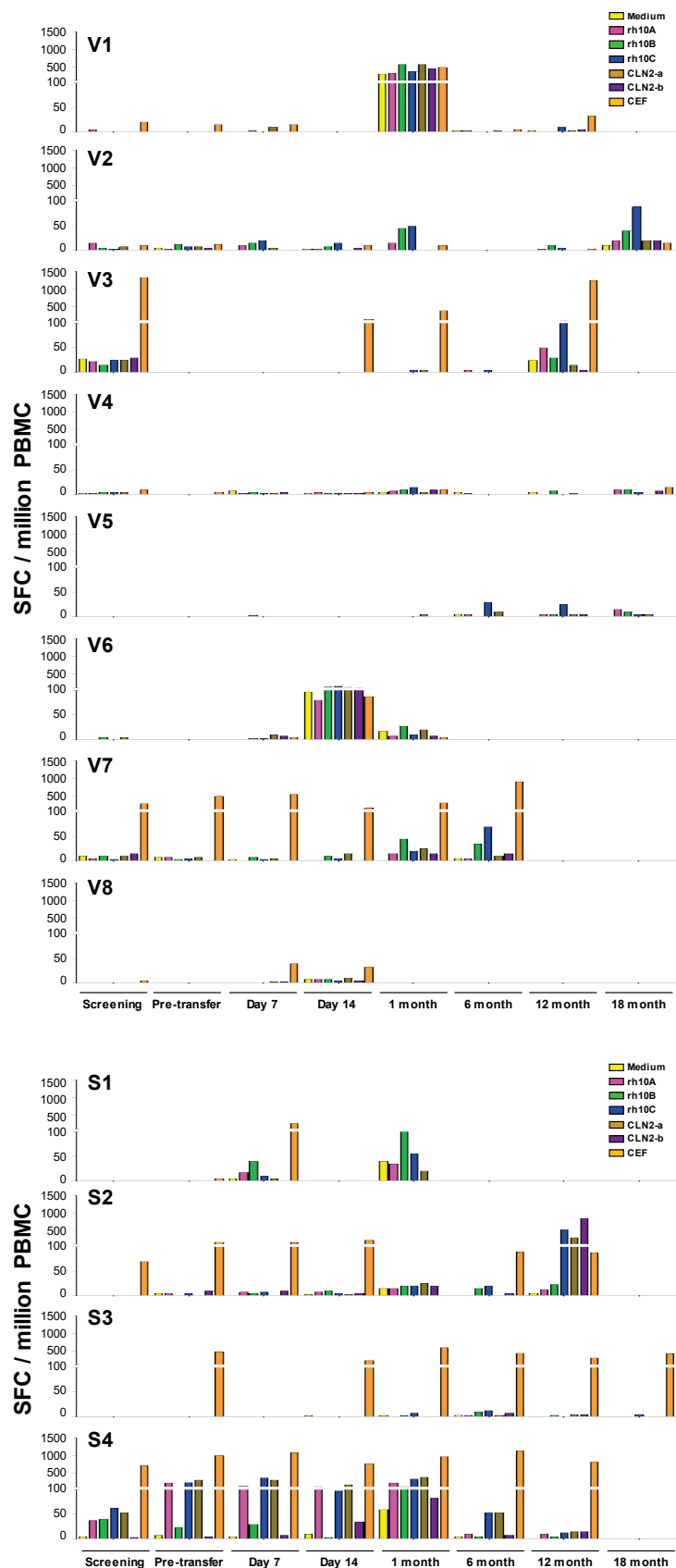
Figure S4. T cell responses to AAVrh.10 capsid and CLN2 transgene. Evaluated by IFN-γ ELISPOT of peripheral blood mononuclear cells (PBMC) stimulated with AAVrh.10 capsid peptides or CLN2 transgene peptides. Stimulation with phytohemagglutinin (PHA) provided the positive assay control; the negative control was growth media. A panel of MHC class 1 restricted viral peptides from cytomegalovirus, Epstein-Barr virus and influenza viruses (CEF) served as a positive control peptide library reference. Data is plotted as spot-forming units per million PBMC. PBMC were derived from sera obtained at 1 or 2 times before (screening and pre-transfer) and at days 7, 14 and months 1, 6, 12 and 18 after vector administration and stimulated with each of 3 pools (A, B, and C) of AAVrh.10 capsid peptides or with each of 2 pools (A and B) of CLN2 transgene peptides or a positive control peptide pool (CEF). Peptide pools were 15-mers overlapping by 10. **A.** Participants in cohort 1. V7 received the lower dose $(2.85 \times 10^{11}$ gc). **B.** Participants in Cohort 4 (samples were not available for S5). SFU = spot forming units. S2 and S3 received the lower dose $(2.85 \times 10^{11}$ gc). In samples from some subjects there was no response to the positive control likely due to these children not having prior exposure to these infectious agents.
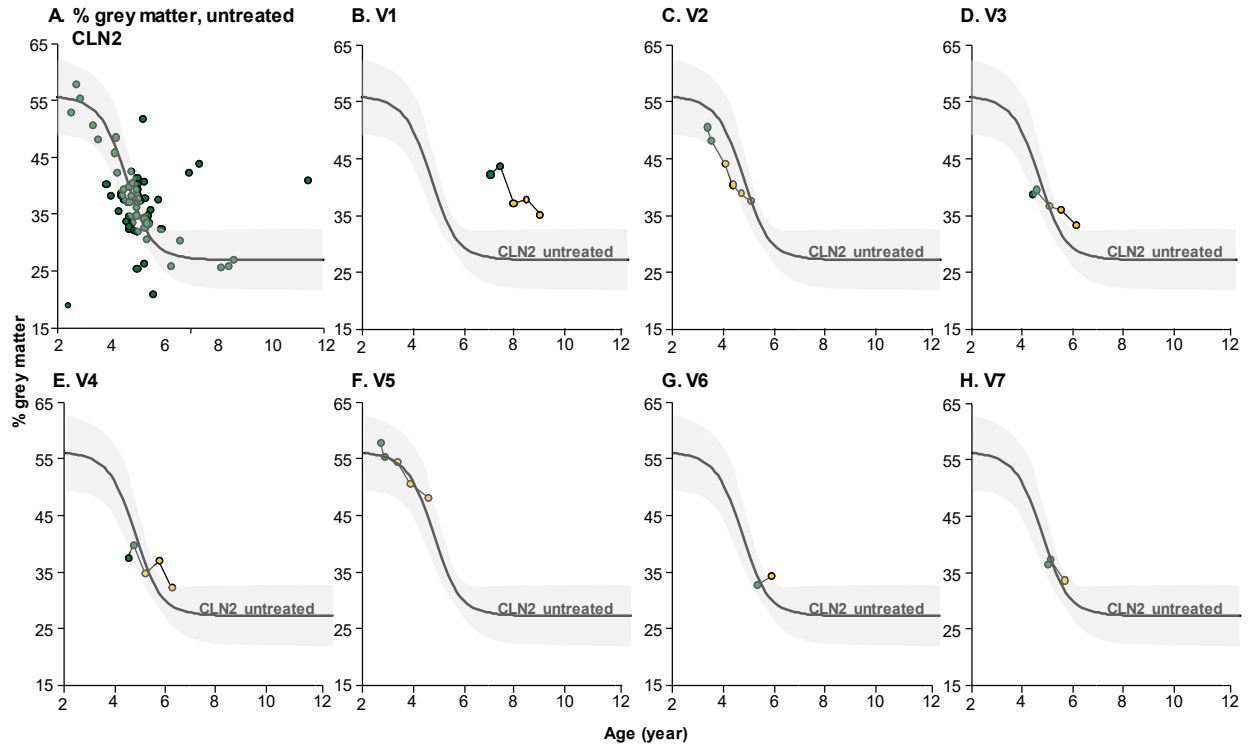
Figure S5. MRI assessment of the treated CLN2 children *vs* untreated CLN2 children. **A.** Percent (%) grey matter *vs* age for untreated CLN2 children (n=62 MRI scans from 47 partcipants, green dots). The solid grey line represents the mean of 1000 bootstrap sigmoidal fits to the CLN2 natural history data. The grey shaded area represents the 95% confidence intervals of those fits. **B-H.** Cohort 1 participants V1 (B), V2 (C), V3 (D), V4 (E), V5 (F), V6 (G), and V7 (H). V7 (panel **H**) received the lower dose (2.85×10^{11} gc). MRI performed on each cohort 1 participant at time-points pre (green dots) and post (tan dots) vector administration overlaid on grey shaded area representing the % grey matter values *vs* age of sigmoidal and linear fits to CLN2 natural history subjects from panel A.
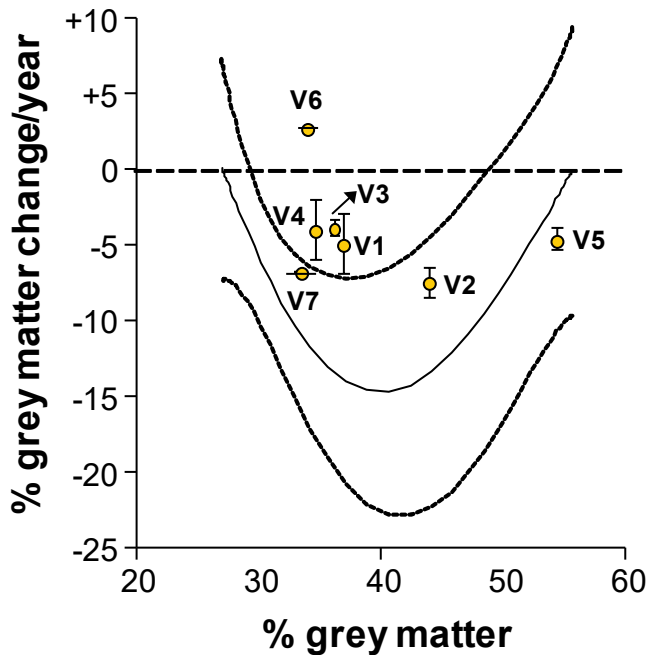
**Figure S6. Quantitative MRI assessment of the treated CLN2 children (cohort 1) *vs*
untreated CLN2 children.** Shown is the MRI % grey matter decline per year vs % grey matter,
as assessed by MRI. In untreated children after birth, the decline in grey matter starts slowly with
near zero slope, then declines rapidly and then slows again, appearing as a sigmoid curve. A plot
of the rate of decline *vs* % grey matter is therefore a parabola with respect to either time or %
grey matter. The dashed line parabolas represent the upper and lower 95% confidence intervals
of the % rate of grey matter decline for the natural history cohort (fig S5A). The solid line
parabola represents the average change of % grey matter change/year as a function of the % grey
matter for the untreated children. Data points from treated subjects with error bars lying above
the upper dashed line had a slower rate of decline of % grey matter compared to untreated
controls. Subject V5 was the youngest trial participant and thus was at the early stage slow rate
of decline and the effect of therapy was not yet apparent. One subject (V6) for whom there was a
slower rate of decline only had one post treatment scan and error bars could not be calculated.
See fig. S5 for the data from which the % grey matter decline/yr of the treated and untreated
children were determined. Subject V7 received the lower dose (2.85×10^{11} gc).
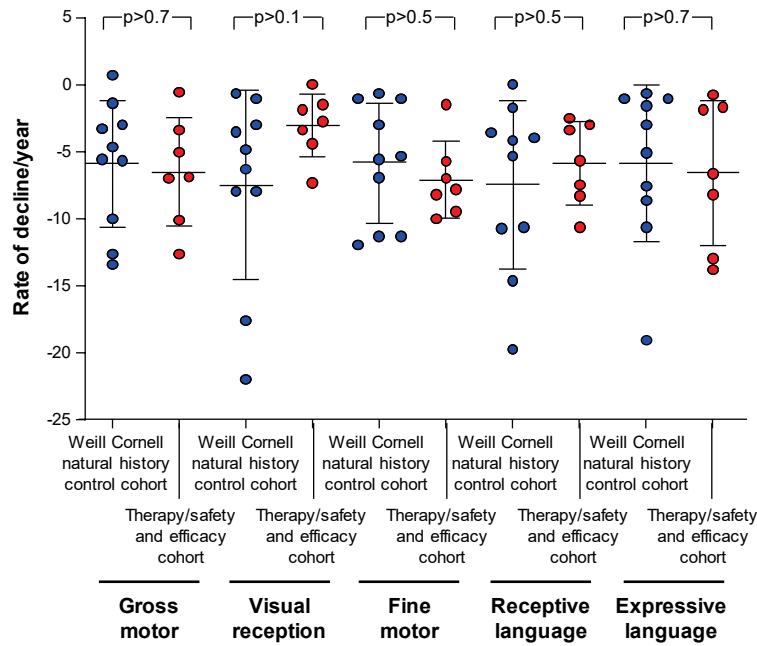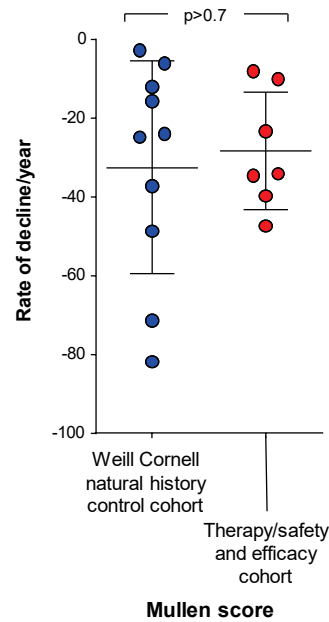
Figure S7. Mullen scale quantitation of the rate of decline for cohorts 1 (red, treated) and 2 (blue, control). A. Linear regression assessed for the scores of each subject in gross motor, visual reception, fine motor, receptive language, and expressive language domains over time to calculate their individual rate of decline within each cohort. **B.** Individual rates of decline for all participants within a cohort were then averaged to calculate the rate of decline/year for the combination of domains or the total Mullen score for each individual cohort. The rates of decline/year for each cohort are plotted as a mean rate of decline with the error bars representing plus and minus one standard deviation from the cohort mean, and the individual rates of declines for each subject are overlaid on the mean data. The p value was calculated using a two-tailed unpaired Student t-test (GraphPad v8.0).
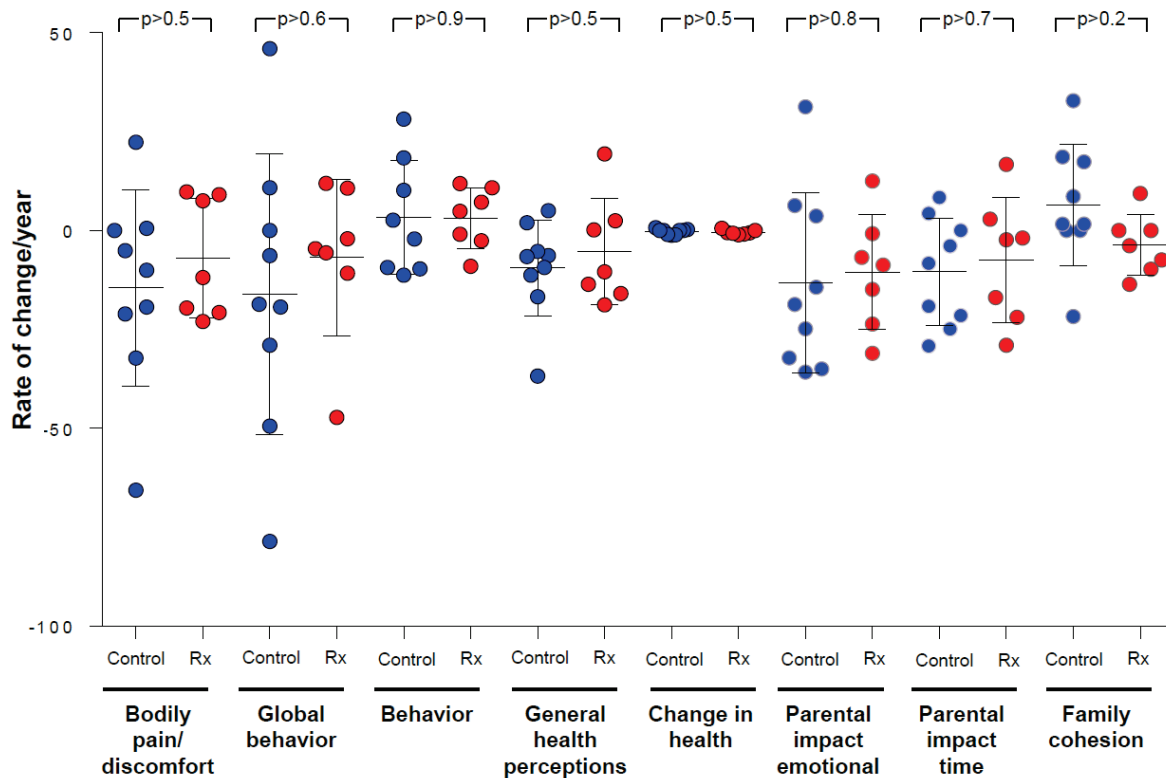
Figure S8. **Impact of treatment on quality of life as assessed by age-dependent quality of life questionnaires.** The parents of all cohort 1 and cohort 2 children were asked to complete either the CHQ or ITQoL (depending on age) quality of life questionnaires. The Infant Toddler Quality of Life Questionnaire Parent Form (ITQoL-PF97, assessing 13 different parameters) was used to evaluate parents of 2 months to 5 year-old participants while the Child Health Questionnaire Parent Form (CHQ-PF50, assessing 14 different parameters) was used to evaluate parents of 5 to 18 year-old participants. During the course of the study as the child aged, they may have aged out of ITQoL and been assessed by CHQ. In order to determine if there was any impact of treatment on the quality of life as determined by these questionnaires, we focused on the 8 parameters that were identical in the two questionnaires. Linear regression assessed for each subjects' scores in each of 8 parameters (bodily pain/discomfort, global behavior, behavior, general health perceptions, change in health, parental impact emotional, parental impact time, family cohesion) over time to calculate their individual rate of decline within each cohort. The rates of decline/year for each cohort are plotted as a mean rate of decline with the error bars representing plus and minus one standard deviation from the mean, and the individual rates of declines for each subject are overlaid on the mean data. The p value was calculated using a two-tailed unpaired Student t-test (GraphPad v8.0). Rx – treated cohort.
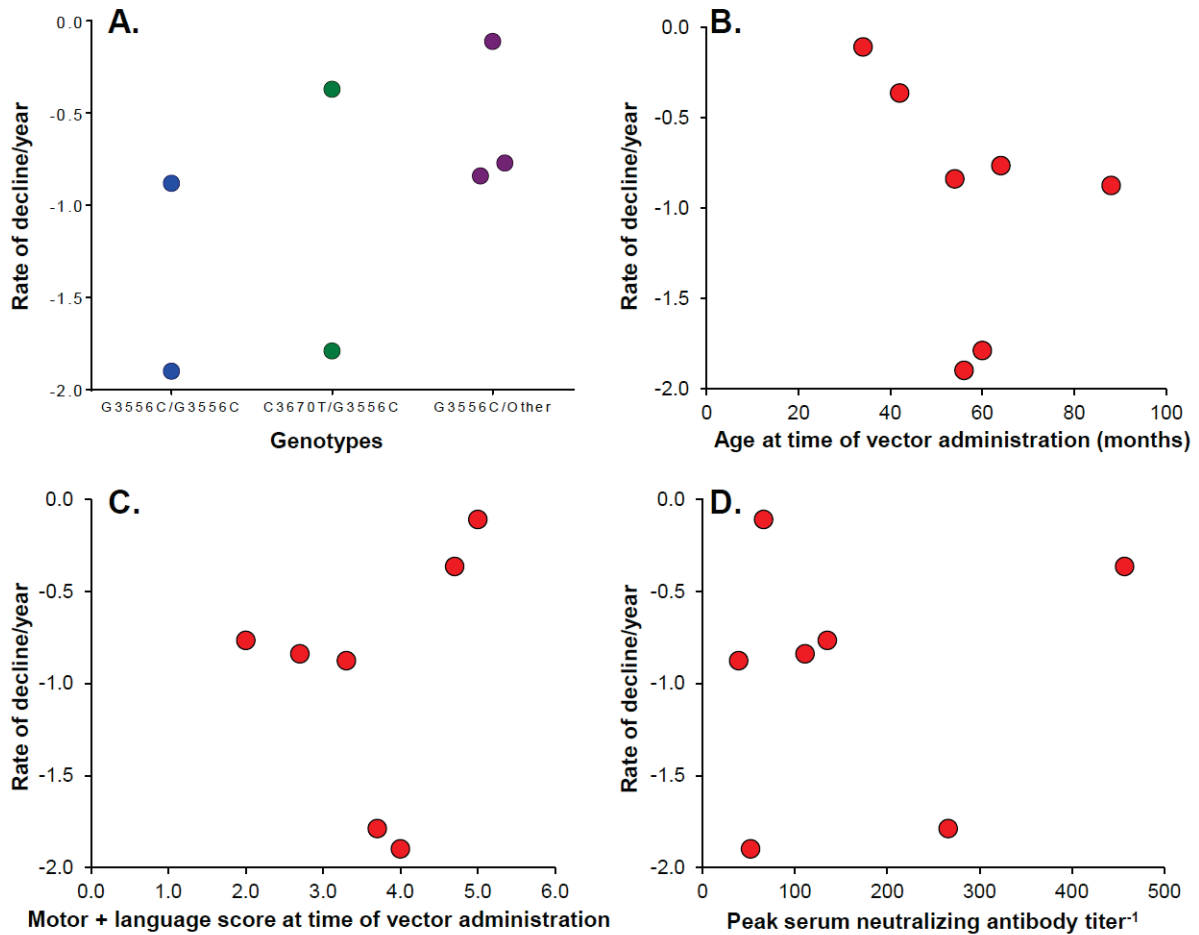
Figure S9. Correlation of various parameters to the rate of decline of motor + language of cohort 1. **A.** Impact of genotype on the rate of decline. Rate of decline for each participant was plotted with respect to genotype: homozygous G3556C/G3556C (blue), heterozygous C3670T/G3556C (green) or heterozygous G3556C/Other genotypes (purple). **B.** Rate of decline as a function of age at time of vector administration. The rates of decline/year for each participant was plotted as a function of age at treatment (in months) for the treated cohort. **C.** Rate of decline as a function of the motor + language score at time of vector administration. The rate of decline/year for each participant was plotted as a function of the combined motor + language scores (scale of 0 to 6) the treated cohort at the time of treatment. **D.** Impact of the peak neutralizing antibody response on the rate of decline. Serum anti-AAVrh.10hCLN2 neutralizing antibody titers were determined at multiple time points through the trial for participants in cohort 1. The rates of decline/year for each subject are plotted against their peak antibody titer.

Table S1. CLN2 Disease Severity Clinical Rating Scales[1]

		Hamburg scale[2]			Weill Cornell scale[3]
Motor	3	Normal	**Gait**	3	Normal
	2	Falls, obvious clumsiness		2	Abnormal, but independent
	1	No unaided walking		1	Abnormal, requires assistance
	0	Immobile		0	Non-ambulatory
Language	3	Normal	**Language**	3	Normal
	2	Abnormal		2	Abnormal
	1	Barely understandable		1	Barely understandable
	0	Unintelligible or no speech		0	Unintelligible or no speech
Visual function	3	Recognizes desirable objects, grabs	**Motor**	3	None of below
	2	Grabbing for objects uncoordinated		2	1 of below
	1	Reacts to light		1	2 of below
	0	No reaction to visual stimuli		0	Myoclonus and chorea / tremor / athetosis and upgoing toes
Seizures	3	None in 3 months	**Feeding**	3	No dysfunction
	2	1-2 seizures per month		2	Mild
	1	1 per month		1	Moderate
	0	>1 per month		0	Gastrostomy tube dependent

Table S2. Vector Infusion Time and Operating Room Surgery and Anesthesia Duration in Cohort 1 Participants

[1] In all children, the total vector volume of 1.8 ml was equally divided among 12 cortical locations delivered through 6 burr holes. There were 2, 150 µl infusions through each burr hole (2 locations at 2 depths through each burr hole), 3 burr holes per hemisphere. The rate of infusion was 2 µl/min. There were slight variations on time based on pump calibration. After the specified dose was administered over a period of ~ 75 min to the 6 sites, the catheters were left in place for 5 min to assure tissue penetration. The catheters were then withdrawn approximately half-way from the bottom of the catheter tract to the brain surface, and the remaining 50% of the dose was administered, in parallel, to each of the 6 sites (the less deep of the 2 sites through the burr hole).

[2] Surgery duration included time from when the surgeon started drilling the burr holes and pre-determined location, to when the last burr hole was sutured.

[3] Duration under anesthesia is a surrogate for the entire time the child was in the operating room.

Patient identifier	Infusion 1 duration[1] (min)	Infusion 2 duration[1] (min)	Total infusion time (min)	Surgery duration[2] (min)	Under anesthesia[3] (min)
V1	76	76	152	362	509
V2	77	75	152	360	472
V3	75	75	150	321	456
V4	75	78	153	355	473
V5	76	75	151	312	411
V6	75	75	150	344	476
V7	77	75	152	380	524
V8	76	75	151	369	494
Average ±SD	75.9±0.8	75.5±1.1	151.4±1.0	350.4±23.5	476.9±34.6

Table S3. Vector Infusion Time and Operating Room Surgery and Anesthesia Duration in Cohort 4 Participants

[1] In subject S2-S5, the total vector volume of 1.8 ml was equally divided among 12 cortical locations delivered through 6 burr holes. There were 2, 150 µl infusions through each burr hole (2 locations at 2 depths through each burr hole), 3 burr holes per hemisphere. The rate of infusion was 2 µl/min. There were slight variations on time based on pump calibration. After the specified dose was administered over a period of ~ 75 min to the 6 sites, the catheters were left in place for 5 min to assure tissue penetration. The catheters were then withdrawn approximately half-way from the bottom of the catheter tract to the brain surface, and the remaining 50% of the dose was administered, in parallel, to each of the 6 sites (the less deep of the 2 sites through the burr hole).

[2] Surgery duration included time from when the surgeon started drilling the burr holes and pre-determined location, to when the last burr hole was sutured.

[3] Duration under anesthesia is a surrogate for the entire time the child was in the operating room.

[4] In subject S1 only, all 12 administrations were carried out in parallel, through 12 catheters. Each burr hole received administration through 2 catheters that were attached together and delivered at different depths.

[5] The average total infusion time was calculated using the data for subjects S2 – S5; S1 was not included in the calculation, see footnote 4.

Patient identifier	Infusion 1 duration[1] (min)	Infusion 2 duration[1] (min)	Total infusion time (min)[5]	Surgery duration[2] (min)	Under anesthesia[3] (min)
S1[4]	77	--	77	342	432
S2	76	76	152	482	630
S3	75	75	150	317	417
S4	75	75	150	362	492
S5	75	75	150	348	468
Average ±SD	75.6 ±0.9	75.3 ±0.5	150.6±0.8	370.2 ±64.6	487.8 ±84.8

Table S4. Cerebral Spinal Fluid Nucleated Cells[1]

[1] Cerebral spinal fluid nucleated cell count
[2] Time after vector administration (months)
[3] V7 received the lower dose (2.85×10^{11} gc).

Cohort 1			Cohort 4		
Subject	**Month[2]**	**Cells/μl**	**Subject**	**Month[2]**	**Cells/μl**
V1	18	0	S2	6	1
V2	12	2		12	0
V3	12	0	S3	6	0
V4	12	0		12	1
V5	12	0		18	1
V7[3]	6	1	S4	6	1
				12	2
				18	0

Table S5. Percent Volume of the Brain with MRI T2 Hyperintensity

[1] No T2 FLAIR was observed and hence volume is listed as 0.0%
[2] nd = scan not done
[3] V7 received a lower dose than V1-V6; see Methods

	Time post-vector administration			
	24 hr	6 month	12 month	18 month
	Volume (%)	Volume (%)	Volume (%)	Volume(%)
V1	0.30	0.24	0.47	0.47
V2	0.00[1]	0.06	0.23	0.20
V3	0.28	0.07	0.15	0.15
V4	0.12	0.11	0.32	0.28
V5	0.21	0.12	0.11	0.06
V6	0.07	0.00[1]	nd[2]	nd[2]
V7[3]	0.11	0.08	nd[2]	nd[2]
Average	**0.16**	**0.10**	**0.26**	**0.23**
Standard deviation	**0.11**	**0.07**	**0.14**	**0.16**

Table S6. Quality of Life Questionaires[1, 2, 3, 4]

[1] The parents of all cohort 1 and cohort 2 children were asked to complete either the CHQ or ITQoL (depending on age) quality of life questionnaires. For each visit, the ✓ indicates the quality of life questionnaire that was completed, and the other is left blank. If for a given visit neither quality of life questionnaire was completed, it states ND (not determined). The Infant Toddler Quality of Life Questionnaire Parent Form (ITQoL-PF97) was used to assess parents of 2 months to 5 year-old subjects while the Child Health Questionnaire Parent Form (CHQ-PF50) was used to assess parents of 5 to 18 year-old subjects

[2] The quality of life questionnaires were completed by at least one parent/legal guardian at the times of assessment. The survey was administered independently to each parent to minimize observer bias if both parents were present

[3] During the course of the study as the child ages, they may age out of ITQoL, and be assessed by CHQ

[4] Not done

[5] This questionnaire was completed at the time of screening (visit 1) and also at the last visit (visit 2, which was typically ≥ 18 months from the screening visit

Subjects	First study visit 1[5]		Last study visit[5]	
	ITQoL	CHQ	ITQoL	CHQ
Cohort 1				
V1		✓		✓
V2	✓			✓
V3	✓			✓
V4	✓			✓
V5	✓		✓	
V6		✓		✓
V7	✓			✓
V8	✓		ND	ND
Cohort 2				
C1	ND	ND	ND	ND
C2	✓			✓
C3		✓		ND
C4	✓		ND	ND
C5	✓		✓	
C6	✓			✓
C7	✓			✓
C8		✓		✓
C9	✓			✓
C10	✓			✓
C11		✓		✓
C12	✓			✓

Table S7. Coefficient of Variation Among Observers in the CLN2 Disease Motor + Language Neurologic Rating Scale

[1] To assess the variance in the measurements of the motor and language domains, we took the data obtained during the screening visit for all subjects

[2] Observer #1 is the "live" observer (pediatric neurologist) who performed the exam for the CLN2 disease neurologic rating scale; observers #2-4 were blinded to any patient or treatment related information and rated the children based on a videotape of the live assessment.

[3] The coefficient of variation (CV; the standard deviation divided by the mean). The CV was calculated for each subject relative to the 4 reviewers to compare the scatter of variables involved in the testing. The average CV is reported for each parameter ± standard deviation of the group

| | Motor at Screening[1] | | | | | | | Language at Screening[1] | | | | | | |
| | Observer[2] | | | | | | | Observer[2] | | | | | | |
Subject	#1	#2	#3	#4	Mean	SD	CV[3]	#1	#2	#3	#4	Mean	SD	CV[3]
V1	1	1	1	1	1.00	0.00	0.00	2	1	2	2	1.75	0.50	0.29
V2	3	2	3	3	2.75	0.50	0.18	2	2	2	2	2.00	0.00	0.00
V3	1	1	1	1	1.00	0.00	0.00	1	1	1	1	1.00	0.00	0.00
V4	2	2	2	2	2.00	0.00	0.00	1	2	2	2	1.75	0.50	0.29
V5	2	2	2	2	2.00	0.00	0.00	1	1	1	1	1.00	0.00	0.00
V6	1	1	1	1	1.00	0.00	0.00	1	1	1	1	1.00	0.00	0.00
V7	3	3	3	3	3.00	0.00	0.00	2	1	2	1	1.50	0.58	0.38
V8	2	2	2	2	2.00	0.00	0.00	2	2	1	1	1.50	0.58	0.38
C1	1	ND	ND	ND	-	-	-	1	ND	ND	ND	-	-	-
C2	1	2	1	1	1.25	0.50	0.40	2	2	2	2	2.00	0.00	0.00
C3	1	1	1	1	1.00	0.00	0.00	1	1	1	1	1.00	0.00	0.00
C4	1	1	1	1	1.00	0.00	0.00	1	1	1	1	1.00	0.00	0.00
C5	3	3	3	3	3.00	0.00	0.00	2	2	2	2	2.00	0.00	0.00
C6	1	1	1	1	1.00	0.00	0.00	2	2	2	1	1.75	0.50	0.29
C7	2	2	1	2	1.75	0.50	0.29	1	2	2	2	1.75	0.50	0.29
C8	3	3	3	3	3.00	0.00	0.00	2	2	2	2	2.00	0.00	0.00
C9	1	2	1	1	1.25	0.50	0.40	1	1	1	1	1.00	0.00	0.00
C10	1	1	1	1	1.00	0.00	0.00	1	1	2	1	1.25	0.50	0.40
C11	1	1	1	1	1.00	0.00	0.00	1	1	2	1	1.25	0.50	0.40
C12	1	1	1	1	1.00	0.00	0.00	2	1	1	1	1.25	0.50	0.40
					Average CV		0.07±0.14					**Average CV**		0.16±0.18

Table S8. Reproducibility of Motor and Language Assessment[1]

[1] Data from n=5 study participants pre-therapy from either cohort 1 or 2, with 3 to 4 observers per data point. The data shown is for repeat assessment on the same child carried out within 42 days, a time when deterioration would not be detectable.

Tests	Motor (M)		Language (L)		Total (M+L)	
Subject	Visit 1	Visit 2	Visit 1	Visit 2	Visit 1	Visit 2
V2	2.7	3.0	2.0	2.0	4.7	5.0
V3	1.0	1.0	1.0	1.7	2.0	2.7
V6	1.0	1.0	1.0	1.0	2.0	2.0
V7	2.0	2.0	1.7	1.7	3.7	3.7
C6	1.0	1.0	2.0	2.0	3.0	3.0

Table S9. Assessments of Motor and Language Parameters for Cohort 2[1]

[1] The motor + language data is provided for all subjects in cohort 2. The clinical assessment of motor + language was performed prospectively using defined standard operating procedures (SOPs) based on 3 to 4 observers, with specific rules on how the data was evaluated. The primary, on-site assessor was a pediatric neurologist who had been trained on implementing the scale. The assessment of each child was videotaped by a trained technician following a SOP for recording the assessment and editing for review by 2 to 3 other pediatric neurologists who were trained on implementing the scale. All were blinded to the subjects' treatment status. In the event of discrepancy of more than 1 point between the 2 blinded scorers, a 3rd pediatric neurologist, also blinded, scored the video in order to act as a tie-breaker. The final score was an average of the assessment of 3 to 4 reviewers (primary + 2 to 3 additional reviewers), minimizing bias and subjective interpretation. The data provided here is the final score.

[2] Subjects C1-C12, Cohort 2, participated in the control arm of the study.

[3] Each subject typically underwent 2 to 3 motor and language assessments.

[4] Motor score – Scale of 0-3, 3 is normal, 2 is abnormal, but independent, 1 is abnormal, requires assistance and 0 is Non-ambulatory

[5] Language – Scale of 0-3, 3 is normal, 2 is abnormal, 1 is barely understandable, requires assistance and 0 is unintelligible or no speech

[6] Composite of motor and language

Subject[2]	Study visit[3]	Age at assessment (months)	Time after first assessment (months)	Motor score[4]	Language score[5]	Total score[6]
C1	1	74.8	0	1.0	1.0	2.0
	2	85.8	+ 11.0	0.0	0.0	0.0
C2	1	55.3	0	1.3	2.0	3.0
	2	60.3	+ 5.0	1.3	1.0	2.3
	3	102.7	+ 47.4	0.0	0.0	0.0
C3	1	65.7	0	1.0	1.0	2.0
	2	70.8	+ 5.1	0.7	0.7	1.4
C4	1	47.8	0	1.0	1.0	2.0
	2	58.1	+ 10.4	0.0	0.3	0.3
C5	1	30.0	0	3.0	2.0	5.0
	2	50.7	+ 20.7	2.0	1.0	3.0
C6	1	51.5	0	1.0	2.0	3.0
	2	52.6	+ 1.1	1.0	2.0	3.0
	3	69.4	+ 18.0	0.0	0.0	0.0
C7	1	57.2	0	1.7	1.7	3.4
	2	60.2	+ 3.0	2.0	1.7	3.7
	3	74.6	+ 17.4	0.0	0.0	0.0
C8	1	62.2	0	3.0	2.0	5.0
	2	80.8	+ 18.6	1.0	1.0	2.0
C9	1	58.9	0	1.3	1.0	2.3
	2	65.5	+ 6.6	1.0	0.0	1.0
	3	81.7	+ 22.8	0.0	0.0	0.0
C10	1	56.4	0	1.0	1.3	2.3
	2	75.1	+ 18.6	0.0	0.0	0.0
C11	1	69.0	0	1.0	1.3	2.3
	2	85.1	+ 16.1	1.0	0.0	1.0
C12	1	59.7	0	1.0	1.3	2.3
	2	74.3	+ 14.6	1.0	0.7	1.7

Table S10. Inclusion/Exclusion Criteria for Cohorts 1 and 2[1]

[1] All individuals who meet the following criteria will be included without bias as to a gender or race/ethnicity. Each case will be individually reviewed with the Eligibility Committee comprised of 3 physicians other than the PI, including a pediatric neurosurgeon, pediatric neurologist and general pediatrician.

[2] Natural history data from 140 genotype-confirmed CLN2 patients from two independent international cohorts (*5*), including our data, were analyzed to provide detailed longitudinal natural history data which demonstrated that the motor and language subscores of the clinical rating scales were an accurate predictor of disease progression and severity. The entire 12-point LINCL scale was used to determine inclusion/exclusion criteria for the study, while the motor + language data only were used to determine efficacy. This is similar to what was done to determine the efficacy of Brineura® (*31*).

Inclusion criteria

- Definitive diagnosis of CLN2 disease, based on clinical phenotype and genotype. The genotype must include at least one of the 5 of the following CLN2 mutant genotypes: C3670T (c.622 C>T, nonsense Arg208 to stop), G3556C (c.509-1G>C, intron 5, splice), G5271C (c.1266 G>C, Gln422His), and G4655A (c.1094G>A, Cys365Tyr). If either parental allele is R447H, the patient was not included in the study. These variants account for a total of 83% of the mutations in the 1999 study by Sleat et al (*1*), 52% in the recent variant compilation by Gardner et al (*11*), and 82% of the mutations in the population screened for the therapy *vs* no therapy study. Our data regarding the natural history of the disease and the studies of Steinfeld et al (*63*), demonstrate that, for these 5 genotypes (genetic constitution), CLN2 subjects have similar clinical course.
- The subject must be between the age of 2 and 18 years
- Subjects will have an average total score of 6 -12 on the Weill-Cornell LINCL scale and the total score should not be outside the 95th percentile confidence limits for age based on Worgall et al (*4*).
- The subject will not previously have participated in a gene therapy or stem cell study.
- Parents of study participants must agree to comply in good faith with the conditions of the study, including attending all of the required baseline and follow-up assessments, and both parents or legal guardians must give consent for their child's participation.
- Sexually active subjects will have to use contraception during the treatment and for 2 months after completion of the treatment.
- If asymptomatic but has one older sibling who has a positive genotype and has clinical manifestations of the disease.

Exclusion criteria:

- Presence of other significant medical or neurological conditions may disqualify the subject from participation in this study, particularly those which would create an unacceptable operative risk or risk to receiving the AAVrh.10hCLN2 vector, e.g., malignancy, congenital heart disease, liver or renal failure
- Subjects without adequate control of seizures to screening, or active enrollment in an investigational medication or device study
- Subjects with heart disease that would be a risk for anesthesia or a history of major risk factors for hemorrhage
- Subjects who cannot participate in MRI studies
- Concurrent participation in any other FDA approved Investigational New Drug
- Subjects with history of prolonged bleeding or abnormal platelet function or taking aspirin
- Renal disease or altered renal function as defined by serum creatinine >1.5 mg/dl at admission
- Abnormal serum sodium, potassium calcium, magnesium, phosphate at grade III or IV by Division of AIDS Toxicity Scale
- Hepatic disease or altered liver function as defined by SGPT >150 U/L, and or total bilirubin >1.3 mg/dL
- Immunosuppression as defined by WBC < 3,000/μL at admission
- Uncorrected coagulopathy during the baseline period defined as INR >1.4; PTT >35 sec; platelets <100,000/mm^3
- Anemia (hemoglobin < 11.0 g/dl at > 2 years of age, with normal serum iron studies)

Table S11. Timeline of the Clinical Study[1, 31]

Footnotes for Supplemental Table III

■= test required □= test optional

1 Parameters listed were mandatory for the study; additional parameters were assessed at the discretion of the physician caring for the individual based on general medical practice for similar neurosurgical procedure in this age group. Cohort 1 and 4 underwent all the tests at the time-points specified above and as specified for specific tests below. Cohort 2, the natural history control cohort (no therapy), underwent the assessments at the screening and the 18 month time-point only.

2 Dose of AAVrh.10hCLN2 administered = 2.85×10^{11}-9.0×10^{11} gc.

3 The "Screening" time was the initial eligibility screening assessment. This assessment was carried out under a "screening/control" protocol. Families of the eligible subjects were given the choice to enter the control group (No Therapy, cohort 2) or gene transfer group (Therapy, cohort 1). The subjects entering the gene transfer group were reassessed within 2 wk pre-transfer. This provides the required baseline safety parameters.

4 The "pre-transfer" studies were carried out within 2 wk of administration of the vector, with the exception of the CNS MRI study which had to be done within 24 hr of administration of the vector. If greater than 2 wk prior to administration of the vector, then all of the parameters were reevaluated (listed as "pre-transfer").

5 The "general assessment" was used to make the diagnosis of LINCL on clinical grounds plus CLN2 genomic analysis; prior genomic analysis was accepted.

6 General – medical history, physical exam, vital signs (blood pressure, heart rate, respiratory rate, temperature).

7 Temperature – Parents/legal guardians measured the temperature of the subject every morning for the first three months post administration of the vector. If the temperature was above 38.5° C (101.3° F), the parents/legal guardians were required to report this to the Department of Genetic Medicine immediately.

8 CBC – complete blood count, included: hematocrit, hemoglobin, white blood count, differential, platelets.

9 ESR – erythrocyte sedimentation rate.

10 Clotting – prothrombin time, partial thromboplastin time.

11 Chemistry – sodium, potassium, chloride, total CO_2, blood urea nitrogen (BUN), glucose, magnesium, uric acid, phosphate, creatinine, alanine amino transferase (SGPT), aspartate amino transferase (SGOT), calcium, serum total protein, albumin, alkaline phosphatase, bilirubin (total).

12 Future (serum) – serum sample frozen for future use.

13 Blood type – necessary prior to the surgical procedure.

14 Urinalysis – appearance, specific gravity, pH, protein, glucose, ketones, bilirubin, number and type of cells, characterization of sediment.

15 Pregnancy test (urine): required for pubescent female.

16 EKG – electrocardiogram. If the subject had a cardiac history, previous EKG results were accepted if within 6 months of surgery or MRI provided it was read by a pediatric cardiologist. If no cardiac history was present, an EKG was not necessary.

17 Level of consciousness, speech, language, cranial nerves, motor strength, motor tone, abnormal movements, reflexes, upper extremity sensation, lower extremity sensations, gait, Romberg test, nystagmus, coordination.

18 Posterior-anterior. Subject's previous chest x-ray was accepted if within 6 months of screening unless there was a significant change in his/her clinical scenario. The month 6 and 12 chest x-rays were optional and were only performed if there was a significant change in the subject's clinical scenario since the previous x-ray.

19 For Cohorts 1 and 4 "Vector-related" studies included assessment of anti-AAVrh.10 neutralizing antibodies, anti-AAVrh.10 cellular response (ELISPOT), and anti-CLN2 cellular response (ELISPOT).

20 For Cohorts 1 and 4 samples of CSF were collected under anesthesia. For safety purposes the CSF was assessed for CSF for routine parameters.

21 Routine ophthalmologic exam; this was carried out to help define the overall status of the LINCL. Anesthesia was administered depending on the ability of the subject to remain still as the doctor performs the tests. The eye exam may have included eye dilation, color photos, electroretinogram (ERG), fluorescein angiography and optical coherence tomography (OCT).

22 Subject's family and/or physician were contacted monthly via telephone 1 month after receiving the vector.

23 For cohorts 1 and 2, a clinical rating scale was administered that included assessment of motor + language parameters of the Weill-Cornell LINCL rating scale. This assessment was videotaped and reviewed by multiple neurologists as described in Methods.

24 For Cohort 1, TPP1 levels were assessed in the CSF at one time-point pre and one time-point post administration.

25 For Cohort 1 and 4, Magnetic Resonance Imaging (MRI) at 0 to 2 days was performed for assessment of safety and clinical post-operative care (exact time was determined at the discretion of the neurosurgeon); For cohorts 1 and 2, MRI studies were carried out to assess grey matter volume on the same scanner (3.0 Tesla).

26 For Cohorts 1 and 2, the Child Health Questionnaire™ (CHQ) or Infant Toddler Quality of Life questionnaire (ITQoL) was administered to at least one parent/legal guardian at the designated visits. The ITQoL was developed for use infants and toddlers ages 2 months to 5 years old. The CHQ is a family of generic quality of life instruments that have been designed and normed for children 5-to-18 years of age.

27 For Cohorts 1 and 2, the subjects were evaluated on the developmental scale and videotaped.

28 The 2 month and 3 month evaluation procedures were performed by the subject's local physician.

29 It was possible to perform the 6, 12 and 18 month evaluation procedures locally at the request of the subject's family. Though it was preferable for the subject and his/her family to return to NYPH-WCMC for the 6, 12 and 18 month follow-up visit, the study team coordinated with the subject's family and/or physician to perform the parameters listed in the timeline of the protocol.

30 For follow-up visits performed off-site: weight, future (serum), ophthalmology and lumbar puncture were optional at the 6, 12 or 18 month visit. CBC, clotting, chemistry and MRI were required at either the 6 or 12 month visit and optional at the 18 month visit. Test values from a recent clinical/ hospital visit was accepted if the study team was unable to obtain the measurements or samples during the off-site visit.

31 The acceptable "time windows" for the assessment days were as follows:

	Pre-vector	Day 7	Day 14	Month 1	Month 6	Month 12	Month 18	Year 2 to annual life time follow up
Screening parameters	8 months to 2 wks pre vector administration							
Pre-transfer (baseline)[a]	2 wk to -1 day prior to the vector administration							
Post vector		± 2 days	± 2 days	± 5 days	± 30 days	± 30 days	± 30 days	± 30 days

[a] Except pre-transfer for the MRI/MRS which must be done within 24 hr of the vector administration

Treatment group — Days: 0, 7, 14; Months: 1–18

Category	Screening[3]	Pre-transfer[4]	0	7	14	1	2[28]	3[28]	4	5	6[29]	7	8	9	10	11	12[29]	13	14	15	16	17	18[29]
AAVrh.10hCLN2 administration[2]			■																				
General assessment for diagnosis[5]	■																						
CLN2 genomic analysis[5]	■																						
Safety parameters																							
General[6]	■	■		■	■	■	■	■			■						■						■
Temperature[7]		■	■	■	■	■	□	□															
Weight[30]	■	■									■						■						■
CBC[8, 30]	■	■			■	■					■						■						■
ESR[9]	■	■																					
Clotting[10, 30]	■	■			■	■					■						■						■
Chemistry[11, 30]	■	■			■	■					■						■						■
Future (serum)[12, 30]	■	■			■	■					■						■						■
Blood type[13]		■																					
Urinalysis[14]	□	■			□	□					□						□						□
Pregnancy test (urine)[15]		■																					
EKG[16]	□	■									□						□						□
Neurological assessment[17]	■	■			■	■					■						■						■
Chest X-ray[18]	□	■									□						□						
Anti-vector and anti-transgene immunity[19]		■		■	■	■					■						■						■
Assessment of cerebral spinal fliud[20, 30]	■	□									□						□						□
Ophthalmology[21]	■	□									□						□						■
Follow up telephone call[22]							□	□	□	□		□	□	□	□	□		□	□	□	□	□	
1° efficacy parameter																							
Motor + language scale[23]	■	■				■					■						■						■
2° efficacy parameters																							
TPP1 levels in CSF[24, 30]	■	□									□						□						□
CNS MRI[25, 30]	■	■	■								■						■						■
CHQ or ITQoL questionnaire[26]	■	□																					■
Mullen Scale[27]	■	□																					■

Table S12. Inclusion/Exclusion Criteria for Cohort 4[1]

[1] All individuals who meet the following criteria will be included without bias as to a gender or race/ethnicity. Each case will be individually reviewed with the Eligibility Committee comprised of 3 physicians other than the PI, including a pediatric neurosurgeon, pediatric neurologist and general pediatrician.

Inclusion criteria
- Definitive diagnosis of CLN2 disease, based on clinical phenotype and genotype. If either parental allele is R447H, the patient was not included in the study. This genotype is associated with a late age at onset and protracted clinical phenotype (*49, 50*). No other genotype restriction.
- The subject must be between the age of 2 and 18 years.
- Subjects will have an average total score of <6 on the Weill-Cornell LINCL scale (*4*).
- The subject will not previously have participated in a gene transfer or stem cell study.
- Parents of study participants must agree to comply in good faith with the conditions of the study, including attending all of the required baseline and follow-up assessments, and both parents or legal guardians must give consent for their child's participation.
- Sexually active subjects will have to use contraception during the treatment and for 2 months after completion of the treatment.
- Parents accept inclusion in the treated safety only group (cohort 4).

Exclusion criteria
- Presence of other significant medical or neurological conditions may disqualify the subject from participation in this study, particularly those which would create an unacceptable operative risk or risk to receiving the AAVrh.10hCLN2 vector, e.g., malignancy, congenital heart disease, liver or renal failure.
- Subjects without adequate control of seizures to screening, or active enrollment in an investigational medication or device study.
- Subjects with heart disease that would be a risk for anesthesia or a history of major risk factors for hemorrhage.
- Subjects who cannot participate in MRI studies.
- Concurrent participation in any other FDA approved Investigational New Drug.
- Subjects with history of prolonged bleeding or abnormal platelet function or taking aspirin.
- Renal disease or altered renal function as defined by serum creatinine >1.5 mg/dl at admission.
- Abnormal serum sodium, potassium calcium, magnesium, phosphate at grade III or IV by Division of AIDS Toxicity Scale.
- Hepatic disease or altered liver function as defined by SGPT >150 U/L, and or total bilirubin >1.3 mg/dL.
- Immunosuppression as defined by WBC <3,000/µL at admission.
- Uncorrected coagulopathy during the baseline period defined as INR >1.4; PTT >35 sec; PLT < 100,000/mm^3.
- Anemia (hemoglobin <11.0 g/dl at >2 years of age, with normal serum iron studies).