THE LANCET

Supplementary appendix

This appendix formed part of the original submission and has been peer reviewed. We post it as supplied by the authors.

Supplement to: Sessa M, Lorioli L, Fumagalli F, et al. Lentiviral haemopoietic stem-cell gene therapy in early-onset metachromatic leukodystrophy: an ad-hoc analysis of a non-randomised, open-label, phase 1/2 trial. *Lancet* 2016; published online June 8. http://dx.doi.org/10.1016/S0140-6736(16)30374-9.

Supplementary Appendixes

Contents

Appendix S1. Skin biopsies: methods and supporting results

Supporting materials and methods

Skin biopsies were performed according to local standards, summarized below. Briefly, skin surface was sterilized with ethanol swabs. Skin punches 3-mm in diameter (Stiefel) were used. The skin punches were pushed perpendicularly to the surface of skin, rotated in a clockwise direction, and advanced until reaching approximately 5 mm depth. A bandage was applied over the wounds. Patients' parents were instructed to change a fresh bandage daily until the wounds have completely healed. This procedure is usually safe and no complications have been encountered to date. Bleeding was normally scarce and only 2 min of compression were required to stop bleeding in most cases. As soon as the dermal cylinders were removed from the skin, they were washed with phosphate-buffered saline (PBS) to remove excess blood. The samples obtained were immediately transferred into 2% glutaraldehyde for 24 hours fixation; tissues were placed in 1% osmium solution for 1.5 h, dehydrated, and embedded in Epon; tissue blocks were sectioned vertically with 1 mm thickness, stained with toluidine blue and analyzed under light microscopy (DM 4000B, Leica). Nerve bundles with myelinated fibres were then identified and the blocks were trimmed and cut into ultrathin sections (about 90 nm); these sections were contrasted with lead citrate and uranyl acetate and examined under an electron microscope (Zeiss EM 900).

Light and electronic microscopic analyses

Evaluation of dermal nerve neuropathological findings was conducted, including loss of myelinated nerve fibers, demyelination, onion bulb formation, axonal degeneration and regeneration, and pathologic deposition. Each feature was observed in both optical and electron microscopy. Myelinated nerve fibers reduction is assessed in a semiquantitative way, as detailed in the table below. The other analyzed parameters were simply assessed as present or not in each studied fascicle. All investigations were performed by a reader blinded to sample identity. Results were then compared.

Supporting table: Comparison of dermal nerve biopsies at baseline and post-transplant

	Baseline					**+2 years after HSC-GT**				
parameters	FL	AS	De	OB	TM	FL	AS	De	OB	TM
MLD01	++	++	++	++	++	+	0	0	0	+
MLD02	0	+	+	0	+	+	0	0	0	+
MLD03	+++	+	++	0	+	+	0	0	0	+
MLD04	0	++	0	0	++	+	+	0	0	+
MLD05	++	+++	++	+	+	+	+	+	+	+
MLD06	++	++	++	+	+	+	+	+	+	+
MLD07	++	+	0	0	+	+	++	++	0	+

Morphological analysis of dermal nerve fibers was carried out to focus on fiber loss, storage material within Schwann cells cytoplasm, demyelination and remyelination. FL: fiber loss; AS: abnormal storage; De: demyelination; OB: onion bulb formation; TM: thin myelin. 0 = no signs; + = mild; ++ = moderate; +++ = severe. Data at +2 years follow up are available up to patient MLD07, thus in the absence of their follow up data, the baseline results for patients MLD08 and 09 are not shown.

Appendix S2. MR imaging detailed methods

MR sequences

All children underwent anatomical MRI on a 1.5 Tesla scanner using a 6-channel SENSE head coil (Gyroscan Intera, Philips, Netherlands) as described[1].

Diffusion tensor images (DTI) were acquired using a SE diffusion EPI sequence with the following parameters: TR = 5.000 ms TE = 80 ms; 15 directions of diffusion gradients, b = 800/1000 s/mm2, SENSE factor = 2.

Spectroscopy H-MR sequences employed are: i) regional single voxel proton MR spectroscopy, PRESS technique, TE 31, volume (20 X 20 X 20 mm) on periventricular postero-superior white matter; ii) multi voxel CSI, 2D Press TE 144, FOV 200, sl h 15 mm, with volume positioned on bilateral centrum semiovale.

Adapted Loes'scale.

Involvement of:		
Supratentorial White Matter	Parieto-occipital	Periventricular
		Central
		Subcortical
	Antero-temporal	Periventricular
		Central
		Subcortical
	Frontal	Periventricular
		Central
		Subcortical
Corpus Callosum	Genu	
	Body	
	Splenium	
Pyramidal Tracts	Internal capsule	
	Brainstem location	
Thalami		
Cerebellum		
Presence of:		
Focal Atrophy	Parieto-occipital	
	Antero-temporal	
	Frontal	
	Genu	
	Splenium	
	Cerebellum	
	Brainstem	
Global Atrophy	Diameter of III ventricle	
	Diameter of the frontal horn of lateral ventricles	
Tigroid Aspect		

The MLD scoring system introduced and described by us[2] was further adapted to increase the sensitivity of the scoring, better quantify even minimal modifications of the white matter signal (pre- or early-symptomatic patients) or the slow progression of atrophy in the late phases of the disease progression (natural history patients). In particular, the previous scoring system attributed a score of 1 for the involvement/presence of each item listed in the table, except for global atrophy which was scored as 1 if the diameter of the third ventricle was < 10 mm, and as 2 if it was >10 mm. The maximum score was 28. The revised scoring is now based on the following:

- extension of white matter involvement: minimum score 0.25 up to a maximum score of 1 for each area (0.25-0.5-0.75-1);
- presence and extension of focal atrophy, range 0.25 -1.5 (discrete points: 0.25);
- presence and extension of global atrophy, range 0.25 – 2.0 (0.25-0.50 if the diameter of the third ventricle is <5mm; 1 if the diameter is 5-10 mm; 2 if the size is >10 mm);
- a maximum score of 1 for the presence of overt tigroid aspect with discrete points of 0.25.

The maximum final score is 31.5.

Two neuroradiologists independently reviewed and re-calculated the scores of all the MRIs performed both by the HSC-GT treated patients and by the patients within the natural history study.

Appendix S3. Statistical analyses: methods and supportive results

With its capability for modeling longitudinal data, the mixed effects model[3] is widely used for the analysis of biological data with repeated measurements. The mixed model has several features: (1) to characterize group and individual behavior patterns in a formal way, (2) to acknowledge both group and individual differences, and (3) to incorporate additional covariates over time (4) to model higher order, nonlinear changes over time.
The mixed model is able to characterize individual behavior through random effect, that is, it naturally represents individual trajectories in a formal way. Therefore, it is a natural choice for analyzing longitudinal data, accounting for the heterogeneity of the subjects.

S3.1 Statistical methods employed for the analysis of longitudinal data
The longitudinal (follow-up from the treatment) trend of the engraftment (expressed in logit scale) was estimated as a function of the parameters related to transduction and to the treatment (total AUC, cell dose, % CD34+ infused, transduction efficiency, VCN CD34+, days of neutropenia, days at 0 neutrophils), by using a linear mixed-effects (LME) model, with a random effect on the intercept to account for the unobserved heterogeneity of the patients that cannot be explained by the covariates. We used the logit transformation of the engraftment to meet the assumption of normality of the model. The LME model analysis was also employed to evaluate the longitudinal (follow-up from the treatment) impact of engraftment on the ARSA activity in CSF. In this model the engraftment was included as a time-varying covariate and the random effect was again set on the intercept.
In the analysis of the MR score and the GMFM score, we tested the differences between the treated and untreated patients over time (age in months), by using a nonlinear mixed-effects (NLME) model, due to the logistic shape of their trajectories over time. Thus, for both the MR score and the GMFM score separately, we used the following logistic model:

$$\frac{Asym}{1+\exp\left[\frac{xmid-age}{scale}\right]},$$

where $Asym$ represents the horizontal asymptote as the age increases (i.e. $Asym$ is the plateau of the response value), $xmid$ represents the inflection point that corresponds to the value of age associated to a response equal to half the value of the asymptote (i.e. $Asym/2$), $scale$ represents the distance between $xmid$ and the value of age at which the response is almost 0.73 time the value of the asymptote (i.e. $0.73 * Asym$). For testing the differences between the groups of patients, we allowed all parameters in the model ($Asym$, $xmid$ and $scale$) to depend on the group. In order to account for the heterogeneity of the patients random effects have been set on intercept of the asymptote $Asym$.
For evaluating the possible dependence of the trend of the NCV index over time (follow-up from the treatment) on the engraftment, we estimated a NLME model with the asymptotic regression model, due to the shape of the longitudinal trend of the data:

$$y = Asym + (R0 - Asym)e^{(-Time*e^{lrc})}.$$

The parameter $Asym$ represents the horizontal asymptote (i.e. the value of the plateau reached by the NCV index). Hence, we studied the possible influence of the engraftment on this parameter, that is whether the heterogeneity of the asymptote values among the patients depends on their engraftment (which was used in the model as a time-varying covariate). The parameter $R0$ represents the intercept, thus we put the random effect on this parameter and we allowed it to be influenced by the NCV index at baseline (i.e. the value of the NCV index of the patient before the treatment). The lrc parameter is the natural logarithm of the rate constant.
For all the LME and NLME model analyses, we used the standard general assumption that the variance-covariance matrix of the random effects is a positive-definite matrix[3]. When necessary, the final models were obtained by using a backward selection procedure on the fixed-effects covariates with removal significant level 0.05. For each analysis, follow-up data with missing values in at least one of the considered variables were excluded from the analysis.

S3.2 Results of the longitudinal statistical analysis

LME model analysis for assessing the dependence of engraftment on parameters related to the transduction and to the treatment

The LME model for the engraftment was estimated by considering as covariates both the time (as follow-up from the treatment) and the parameters related to the transduction and to the treatment (as detailed in Section S3.1). The final model, obtained with the backward selection procedure, is:

$$logit(engraftment) = \beta_0 + \beta_1(VCN\ CD34+) + \beta_2(Days\ at\ 0\ neutrophils)$$

parameter	estimate	p-value
β_0	-3.21	0.0008
β_1	0.61	0.0375
β_2	0.30	0.0221

From the results of the estimated model, the longitudinal trend of the engraftment does not depend on the time of follow-up (the coefficient of the time resulted to be not significant), hence the engrafment appears to be constant over time after initial fluctuations (in the first 3 months post-transplant). The only parameters that significantly (and positively) affected the engraftment are: the days of absolute neutropenia and the VCN of the medicinal product.

LME model analysis for assessing the impact of the engraftment on ARSA activity measured in the CSF

As described in detail in Section S3.1, the starting longitudinal model for the ARSA activity in CSF was:

$$ARSA\ activity = \beta_0 + \beta_1(Time) + \beta_2(engraftment).$$

parameter	estimate	p-value
β_0	0.892	0.0333
β_1	0.003	0.7097
β_2	0.453	0.5318

No covariates are retained in the model, after applying the backward selection procedure. Therefore, ARSA activity in CSF does not seem to change over time and is independent of the engraftment.

NLME model analysis for evaluating the difference in MR score between treated and untreated LI patients

Starting from a full NLME model where all the parameters depended on the group (as described in Section S3.1), we obtained the following final longitudinal model for the MR score, with the backward selection procedure:

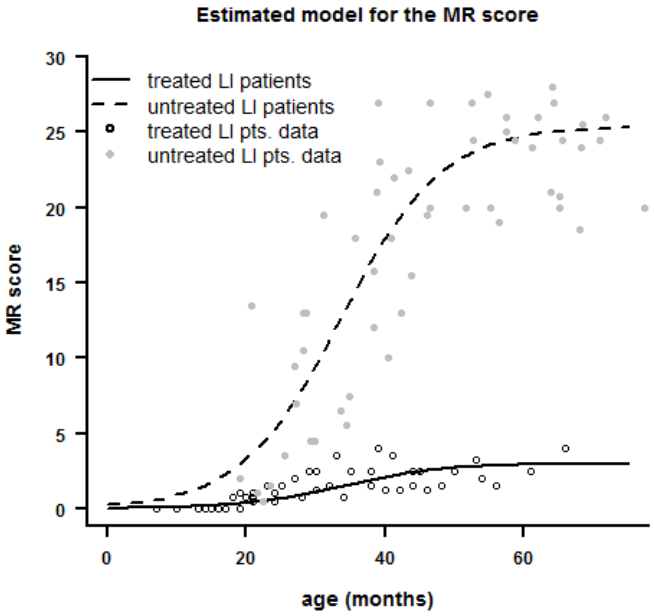
$$\text{MR score} = \frac{Asym}{1 + \exp\left[\frac{xmid - age}{scale}\right]},$$

with

$$Asym = \beta_0 + \beta_1(Untreated)$$

where $Untreated = 1$, for the untreated LI patients, and otherwise $Untreated = 0$.

parameter	estimate	p-value
β_0	3.01	0.0800
β_1	22.39	<0.0001
$xmid$	33.68	<0.0001
$scale$	7.21	<0.0001



The estimated model shows that untreated patients have a significant higher plateau of the MR score with respect to treated patients. The estimated difference in the plateau level of the MR score between the two groups is 22.39.

NLME model analysis for assessing the impact of the engraftment on the NCV index
As described in detailed in Section S3.1, the longitudinal model of the NCV index in dependence on the engraftment is:

$$y = Asym + (R0 - Asym)e^{(-Time*e^{lrc})},$$

where
$Asym = \beta_0 + \beta_1(engraftment)$, $R0 = \gamma_0 + \gamma_1(NCV\ index\ baseline)$.

parameter	estimate	p-value
β_0	-16.20	0.0003
β_1	13.72	0.0255
γ_0	0.41	0.8344
γ_1	1.07	0.0002
lrc	-2.98	<0.0001

The engraftment shows a positive significant effect on the plateau reached by the NCV index (p-value = 0.0255). As expected, the intercept of the curve depends on the value of the NCV index at baseline (p-value = 0.0002).

NLME model analysis for evaluating the difference in the GMFM score among healthy controls, and treated subjects and untreated LI patients
When considering all healthy controls and patients, the estimated longitudinal model for the GMFM score is:

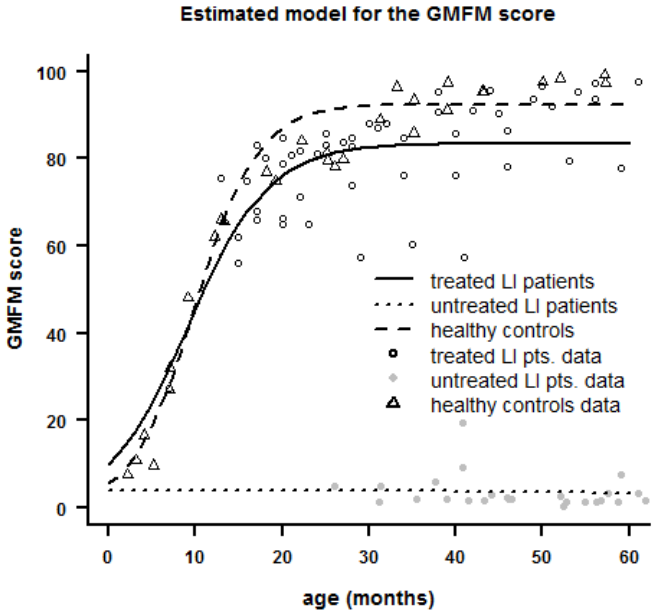
$$\text{GMFM score} = \frac{Asym}{1 + \exp\left[\frac{xmid - age}{scale}\right]},$$

with

$$Asym = \beta_0 + \beta_1(HC) + \beta_2(Untreated), xmid = \gamma_0 + \gamma_1(HC) + \gamma_2(Untreated),$$
$$scale = \delta_0 + \delta_1(HC) + \delta_2(Untreated),$$

where $Untreated = 1$, for the untreated LI patients, and otherwise $Untreated = 0$, and $HC = 1$, for healthy controls, and otherwise $HC = 0$.

parameter	estimate	p-value
β_0	83.43	<0.0001
β_1	9.03	0.0030
β_2	-79.40	<0.0001
γ_0	9.23	<0.0001
γ_1	0.79	0.2586
γ_2	78.23	0.0100
δ_0	4.62	<0.0001
δ_1	-1.01	0.0681
δ_2	-22.66	0.2942



The plateau of the GMFM score of treated patients is significantly greater than the one of untreated LI patients (p-value <0.0001). Although, the plateau of the GMFM score of treated patients is also significantly lower from the one of healthy controls (p-value = 0.0030), it is much closer to the latter (estimated difference = 9.03) than to the one of the untreated LI patients (estimated difference = 79.40). When considering only healthy controls and those patients treated when really devoid of any disease manifestation (thus excluding patients MLD01 and MLD07, as discussed in the text), the estimated longitudinal model for the GMFM score is:

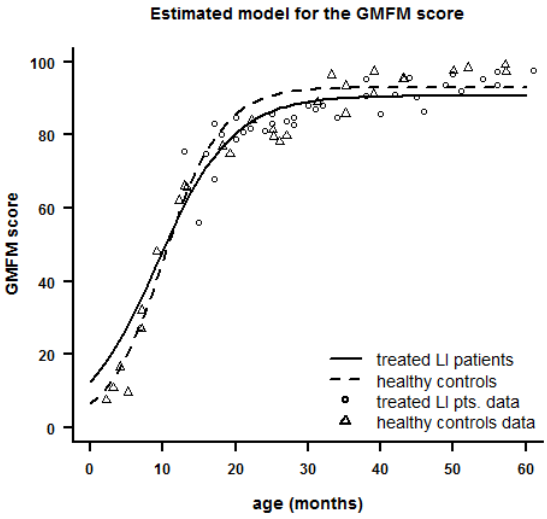
$$\text{GMFM score} = \frac{Asym}{1 + \exp\left[\frac{xmid - age}{scale}\right]},$$

with

$$Asym = \beta_0 + \beta_1(HC), xmid = \gamma_0 + \gamma_1(HC), scale = \delta_0 + \delta_1(HC)$$

where $HC = 1$, for healthy controls, and otherwise $HC = 0$.

parameter	estimate	p-value
β_0	90.65	<0.0001
β_1	2.33	0.4321
γ_0	9.37	<0.0001
γ_1	0.85	0.3537
δ_0	5.14	<0.0001
δ_1	-1.19	0.1298



In this case, excluding patients MLD01 and MLD07, healthy controls and treated patients do not show anymore a significant difference in the plateau of the GMFM score (p=0.4321).

Appendix S4. MLD09 classification

Patient MLD09 was diagnosed as affected by pre-symptomatic early onset MLD due to the reported disease onset in her older sibling between 24 and 36 months of age; due to an absence of patient medical records and the length of time elapsed from initial symptom onset in the older sibling to the time the patient presented for gene therapy study consideration, a more precise estimation of clinical onset was not possible. The patient and her older sibling are homozygous for a mutation (G309S) previously found both in LI patients in association with a severe allele and in Adult MLD subjects in association with a mild mutation[4-6]. This mutant allele showed residual enzyme activity (13% of normal values) and sulfatide degrading capacity (around 50% of normal).[6] These findings, together with the course of the older affected sibling (alive at age 16) that appears much milder than the typical LI course and similar to the typical EJ course (Figure 4B), and the phenotype of MLD09 (i.e. presence of a severe peripheral neuropathy and lack of overt cerebral demyelination in the pre-symptomatic stage) suggest that MLD09 and her older sibling could be considered as patients affected by a clinical variant of intermediate severity between the classical LI and EJ forms.

Appendix S5. MLD04 clinical course

S5.1 Clinical course description

MLD04 showed severe demyelination and motor and cognitive dysfunction at enrollment and baseline; moreover, the patient experienced rapid disease progression between enrollment, treatment and the early post-treatment follow up, as documented by clinical and instrumental assessments described in the text below and Figure S1. Of note, the patient's pre-transplant evaluation phase was slightly longer than in the other patients (2 months) due to the need to collect mobilized peripheral blood to enable manufacture of an adequate medicinal product dose, and the patient experienced an episode of severe metabolic acidosis post-transplant, as recently described. [7]

The patient had an unremarkable developmental history, with normal developmental milestones. She started walking before 18 months.

At *disease onset*, when MLD04 was 4 years and 6 months old, motor abnormalities (more pronounced in fine movements than gross movements) were reported. Two months later, at time of *diagnosis*, neurological evaluation revealed drooling, monotonous and slow speech; decreased symmetric muscle tone; bilateral positive Babinski sign; independent gait, with a wide base, ataxic and periodically stumbling; significant difficulty in climbing on a chair. The GMFC-MLD was equal to 1. Neuropsychological evaluation demonstrated an intellectual development within normal limits (Columbia Scale, after conversion to Wechsler Scale, IQ of 86). Brain MRI showed in the white matter of both hemispheres, in frontal lobe, parietal lobe and on both sides in the fronto-parieto-temporal borders, extensive symmetric hyperintense signal in FLAIR and T2 images. ENG recording documented a sensory and motor demyelinating polyneuropathy.

At *screening for enrollment* into the HSC-GT trial, when the patient was 4 years 9 months old, the neurologic evaluation revealed the presence of hypoactive deep tendon reflexes (DTRs) in the upper limbs, very brisk DTR at the lower limbs, and Achilles clonus (new findings). The GMFC-MLD was equal to 1.

At *baseline*, when MLD04 was 4 years and 10 months old, the neurologic evaluation revealed the presence of bilateral equinus-cavus foot. The GMFC-MLD was equal to 2. GMFM was 73.9%. Neuropsychological evaluation revealed a decline in IQ as compared to the previous assessment, with a value of 58 (Wechsler Preschool and Primary Scale of Intelligence). Brain MRI was comparable to the previous one, with a total score of 11. ENG confirmed the sensory and motor demyelinating polyneuropathy with a NCV score of -8.26.

At *1 month post-treatment*, when the patient was 5 years old, the neurological evaluation revealed a marked worsening of the clinical condition: the exam showed MLD04 was able to control her head but with a tendency to fall backwards, kyphosis of the trunk, bilateral support required to maintain standing position, inability to walk independently; reduced spontaneous motility of the four limbs, tendency to maintain lower limbs externally rotated, and reducible bilateral equinus varus-supinated feet. GMFC-MLD worsened to 3.

At *+ 3 months post-treatment* further worsening of the clinical condition was noted at the neurological assessment: the patient was unable to manipulate objects, could pronounce only syllables; upper limbs and fingers were kept in flexion with mild spastic hypertonia (lower limbs in hyperextension with bilateral equinus varus-supinated feet reducible to 90°); DTRs were reduced/absent at four limbs; ataxia was present at upper limbs, head and trunk. The patient was able to maintain head on the midline with a tendency to fall forward and to control the trunk in the sitting position for more than 10 seconds, with kyphosis. She could reach and maintain the standing position with bilateral support. She could take a few steps on tiptoe with bilateral support. GMFC-MLD was 3, GMFM was 22.4%. Brain MRI showed a progression of the neuroradiological findings, with increased size of the ventricular system and subarachnoid spaces, extension of the known widespread alteration of the signal of the white matter in the anterior and central areas, and reduction of the peak of N-acetyl-aspartate and the presence of labeled myoinositol peak evident at short echo time at spectroscopy. NCV index was stable but with reduced amplitude of motor action potential of right ulnar and deep peroneal nerves and of sensory action potential of right median nerve.

At *+ 6 months post-treatment* further deterioration was noticed at neurological assessment: the patient was no longer able to pronounce syllables and to carry out orders; the bilateral equinus varus-supinated feet became difficult to reduce to 90°; lower limbs DTRs evoked clonus bilaterally; the patient could maintain head on the midline for only few seconds with tendency to fall forward. GMFC-MLD was 4, GMFM 17.6%. Neuropsychological assessment revealed further worsening with an IQ < 40. ENG compared with the previous evaluation showed mild improvement of latency and amplitude of motor

action potential of deep peroneal nerve, with stable NCV index. Neuroradiological findings were stable.

From the + 6 months post-treatment follow up a trends towards progressive stabilization of the patient's clinical condition was observed, as documented also in Figure S1. A tendency to mild improvement in the NCV index was observed. Noticeably, the patient did not develop seizures and other complications typically associated with early onset MLD, such as severe dysphagia and aspiration pneumonia, and had preserved some interaction with the examiners and caregivers, with adequate reactions to stimuli.

The clinical experience with this patient (and with one recently treated LI patient – data not shown) resulted in protocol amendments to exclude patients with rapidly progressing clinical symptomatology observed between enrollment and treatment initiation.

Appendix S6. LV IS Analysis

Integration (insertion) site (IS) analysis has been shown to be useful as a surveillance tool in several clinical gene therapy trials [8].

To retrieve and map the vector IS present in HSC gene therapy patients' blood or bone marrow (BM) – derived cells, we used linear-amplification-mediated polymerase chain reaction (LAM-PCR), in combination with high-throughput sequencing and bioinformatics analysis [9,10] Given the semirandom nature of vector integration, each tranduced cell (and its progeny) will harbor a distinctive genetic mark represented by a vector integration in a specific genomic position. The number of IS retrieved from a cell population is proportional to the the extent of transduction (average number of integrated vector copies per cell) and the number of vector marked clones present in the whole population, enabling clonalilty and clonal abundance evalutations. [8]

IS analyses described in this section were performed on genomic DNA extracted from peripheral blood (PB) and BM cells harvested at different points after therapy from the first 7 MLD gene therapy patients (6 LI and 1 EJ). The overall follow up time varies from 18 to 48 months. All procedures were performed as previously described. [9,10] Relevant numbers of integration sites were obtained (ranging from ~2,000 to 9,000 IS in each patient, see the following Table) [11] in 6 MLD patients, indicating polyclonal reconstitution. MLD05 yielded only 237 unique IS, indicating relatively oligoclonal reconstitution, consistent with the lower VCN values in this patient.

Patient	Follow up [months]	Number of LAM-PCRs	Number of Raw Reads	Number of Reads of IS	Number of IS
MLD01	48	110	11,380,055	3,282,525	9,174
MLD02	42	111	17,480,288	4,836,352	7,304
MLD03	42	101	20,218,007	6,118,539	9,189
MLD04	30	57	7,363,206	4,981,275	4,305
MLD05	24	41	6,722,218	1,326,943	237
MLD06	18	66	9,794,604	1,788,616	1,973
MLD07	18	63	9,854,798	2,317,547	2,493
Totals		*549*	82,813,176	24,651,797	34,675

Summary Table reporting for each patient the number of LAM-PCR processed per patient, the number of sequencing reads obtained by next generation sequencing from the LAM-PCR samples, and the number of IS (both by reads, column "Number of Reads of IS" corresponding to the overall sequence count, and by distinct genomic positions, column "Number of IS").

Overall, no sustained clonal dominance was observed in any of MLD patients. Some IS sporadically showed an increase in the abundance at a single time point and this mainly occurred during the early phases of hematopoietic reconstitution.

Common Insertion Sites (CIS) analysis was performed as an additional measure to estimate clonal proliferation in the MLD patients using the Grubbs test for outliers. [12] CIS analyses provide relevant information on the overall genotoxic risk of a vector throughout the study population but may not be used to predict oncogenesis for any given individual patient. CIS are genomic regions targeted by vector integrations at a frequency significantly higher than expected with respect to a random distribution. All analyzed patients shared several CIS that were clustered within megabase-wide genomic regions and targeted genes such as *KDM2A*, *PACS1*, HLA genes and many others. These regions are the same as those described in the previous MLD IS analyses, other LV based gene therapy trials, and in nonclinical studies with LV-marked CD34+ cells. [9,10] On the other hand, no overlap was found with the CIS associated with clonal dominance and/or leukaemic development in the some γ-RV clinical trials. Overall, these data suggest that the CIS observed in MLD are the product of an intrinsic integration bias of this LV vector rather than a hallmark of insertional mutagenesis and abnormal clonal proliferation. Over-representation analysis of gene classes targeted by LV integrations in MLD patients performed using the GREAT software [http://bejerano.stanford.edu/great] showed over-representation of genes involved in immune regulation and chromatin remodeling, which have been also found to be over-represented in the LV adrenoleukodystrophy clinical trial, and no preference to target oncogenes or genes associated with cell proliferation were detected. [9,10]

In summary, IS analysis of the first 7 MLD patients shows a polyclonal pattern of integrations for up to 48 months. Moreover, there was no evidence of stable clonal dominance or evidence of skewing towards gene classes involved in cancer.

REFRENCES

1. Biffi A, Montini E, Lorioli L, et al. Lentiviral hematopoietic stem cell gene therapy benefits metachromatic leukodystrophy. Science 2013;341:1233158.
2. Biffi A, Cesani M, Fumagalli F, et al. Metachromatic leukodystrophy - mutation analysys provides further evidence of genotype-phenotype correlation. Clinical Genetics 2008;74:349-57.
3. Pinheiro JC, Bates DM. Mixed-Effects Models in S and S-PLUS. New York: Springer-Verlag; 2000.
4. Berna L, Gieselmann V, Poupetova H, Hrebicek M, Elleder M, Ledvinova J. Novel mutations associated with metachromatic leukodystrophy: phenotype and expression studies in nine Czech and Slovak patients. Am J Med Genet 2004;129A:277-81.
5. Luzi P, Rafi MA, Zaka M, et al. Biochemical and pathological evaluation of long-lived mice with globoid cell leukodystrophy after bone marrow transplantation. Molecular genetics and metabolism 2005;86:150-9.
6. Kreysing J, Bohne W, Bosenberg C, et al. High residual arylsulfatase A (ARSA) activity in a patient with late-infantile metachromatic leukodystrophy. American journal of human genetics 1993;53:339-46.
7. Lorioli L, Cicalese MP, Silvani P, et al. Abnormalities of acid-base balance and predisposition to metabolic acidosis in Metachromatic Leukodystrophy patients. Molecular genetics and metabolism 2015.
8. Brugman MH, Suerth JD, Rothe M, et al. Evaluating a ligation-mediated PCR and pyrosequencing method for the detection of clonal contribution in polyclonal retrovirally transduced samples. Human gene therapy methods 2013;24:68-79.
9. Aiuti A, Biasco L, Scaramuzza S, et al. Lentiviral hematopoietic stem cell gene therapy in patients with Wiskott-Aldrich syndrome. Science 2013;341:1233151.
10. Biffi A, Montini E, Lorioli L, et al. Lentiviral hematopoietic stem cell gene therapy benefits metachromatic leukodystrophy. Science 2013;341:1233158.
11. Calabria A, Leo S, Benedicenti F, et al. VISPA: a computational pipeline for the identification and analysis of genomic vector integration sites. Genome medicine 2014;6:67.
12. Biffi A, Bartolomae CC, Cesana D, et al. Lentiviral vector common integration sites in preclinical models and a clinical trial reflect a benign integration bias and not oncogenic selection. Blood 2011;117:5332-9.

Figure S1

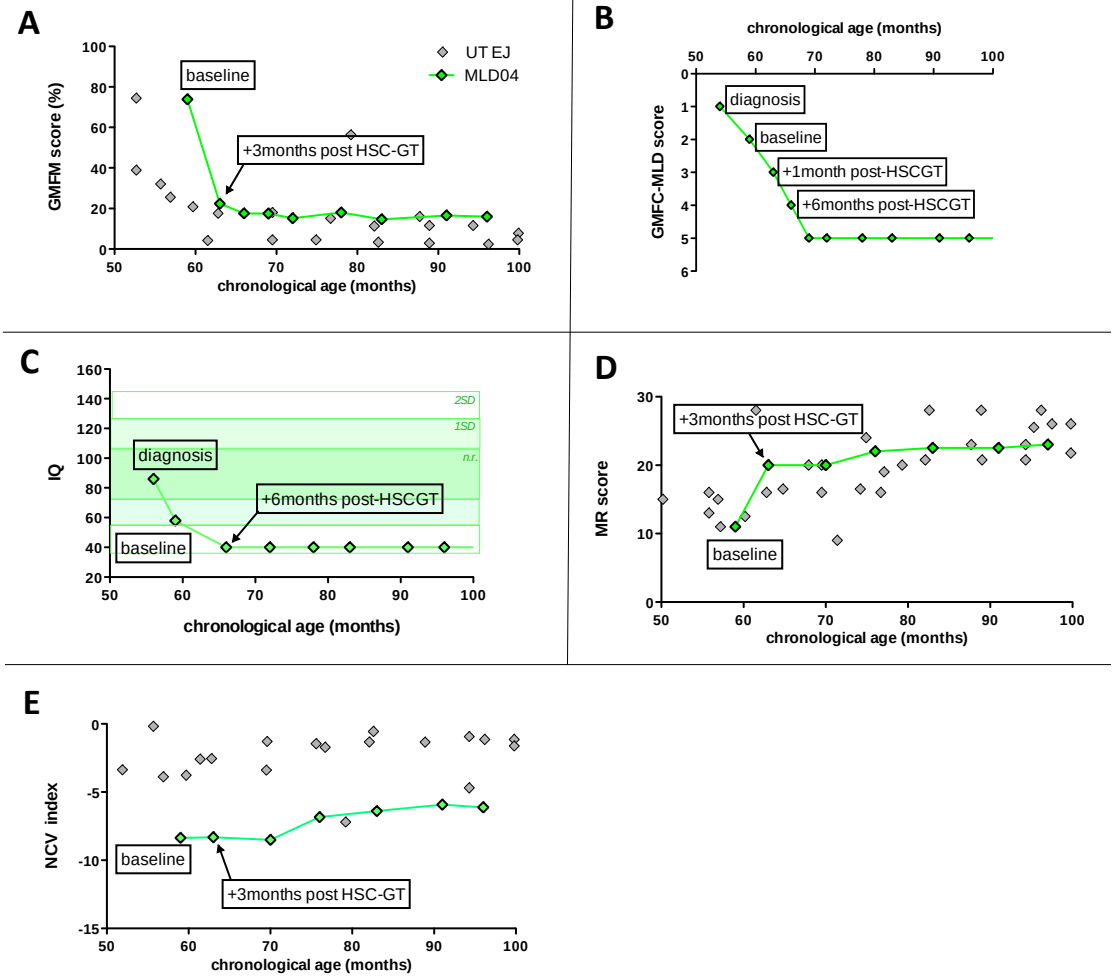


Figure S1. Clinical and instrumental follow up of MLD04. GMFM score (A), GMFC-MLD score (B), total IQ assessed by WPPSI-III scale (B), MR score (D) and NCV index (E) of MLD04 (green diamond) and of a historical cohort of untreated EJ-MLD (grey diamonds) patients (EJ MLD are not shown for GMFC-MLD). The time at each assessment is highlighted in the text boxes. MLD04 experienced a rapid motor and cognitive decline in the phases between diagnosis and early post-HSCGT follow up. The NCV index is the only parameter that slightly improved and stabilized after treatment, starting from 1 year post HSC-GT.

Figure S2

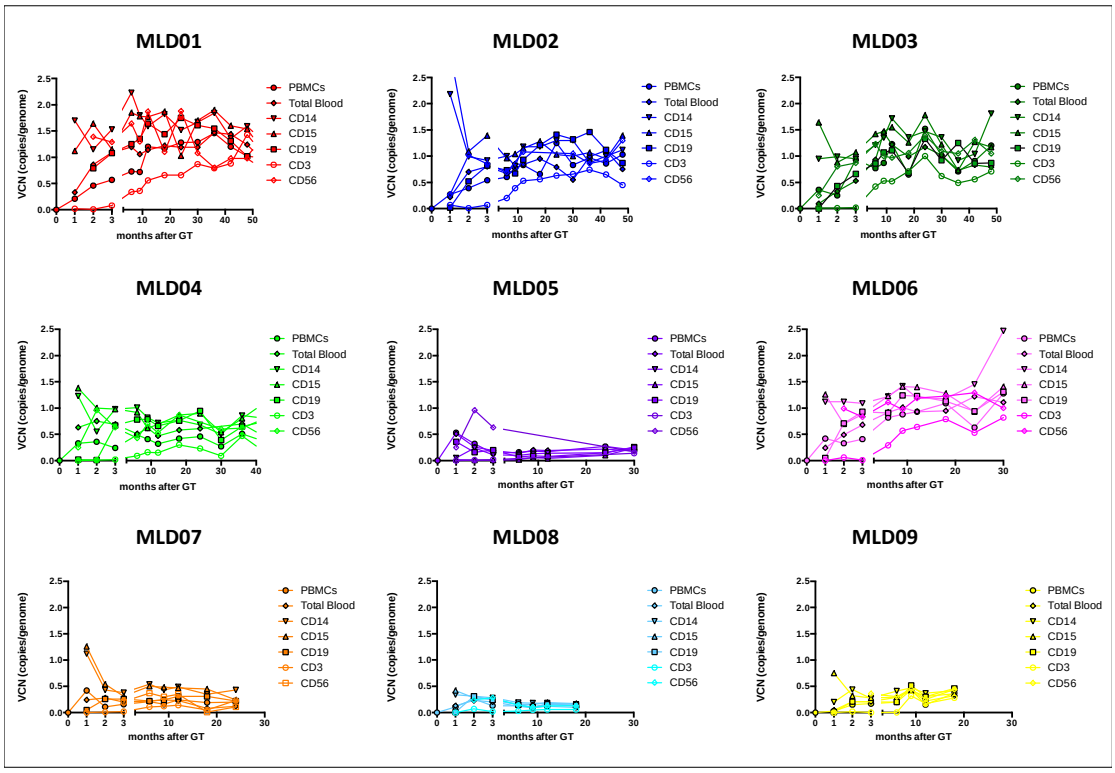


Figure S2. VCN on PBMC subpopulations. Vector copy number (VCN) values measured on PBMCs, total blood and their subpopulations (see graph legend, CD14: monocytes, CD15: granulocytes, CD19: B lymphocytes, CD3: T lymphocytes, CD56: NK cells) in HSC-GT patients during the post-treatment follow up.

Figure S3

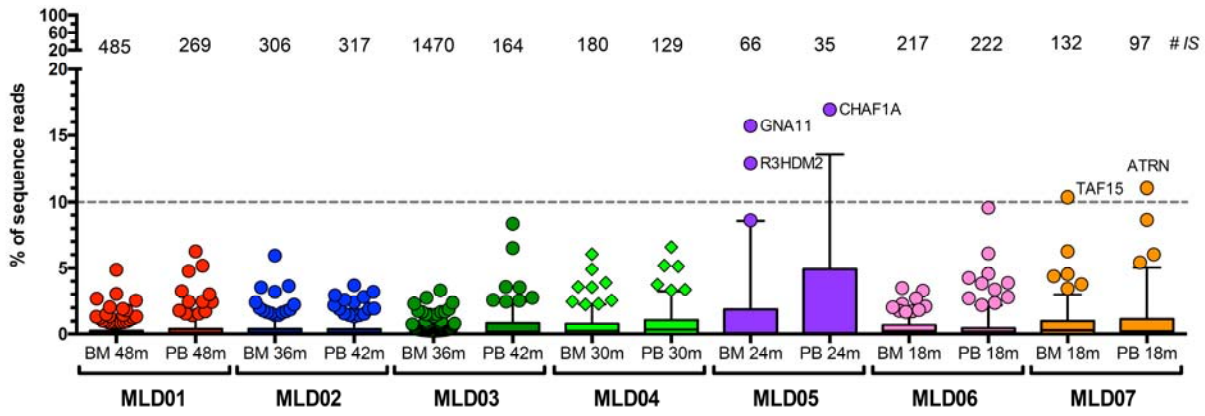


Figure S3. Per patient box-plot representation of the vector integration site (IS). Per patient box-plot representation of the vector integration site (IS) abundance, computed as the percentage of sequence reads from different sources (BM or PB), at the latest available time point for patients from MLD01 to MLD07. The total number of unique IS retrieved from samples (IS#) is reported on top of the graph. ISs outside the 5-95 percentile range are shown as dots (the bottom range is mostly flattened on this plot). Most abundant integrations are reported with the nearest hit gene labeled next to the dot if overcoming the threshold of 10%.

Figure S4

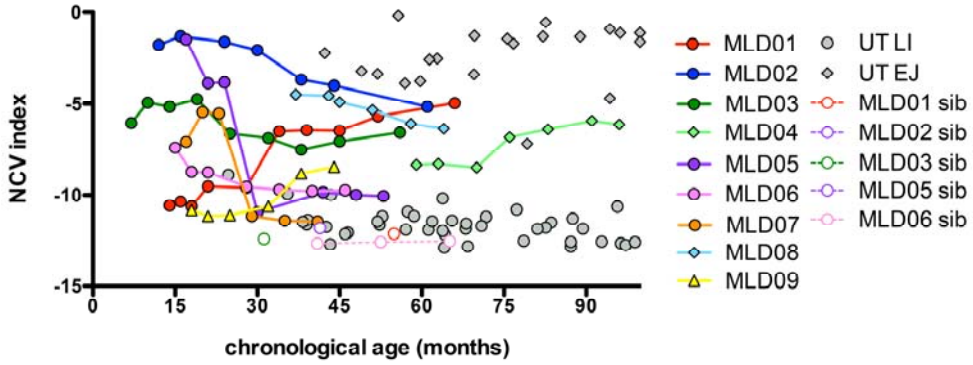


Figure S4. NCV index of early onset treated patients. NCV index of the early onset patients treated when pre- or early symptomatic (see symbol and color-code), of respective older affected siblings (see color-code and dotted lines) and of a historical cohort of un-treated LI-MLD (grey circles) and EJ-MLD (grey diamonds) patients.

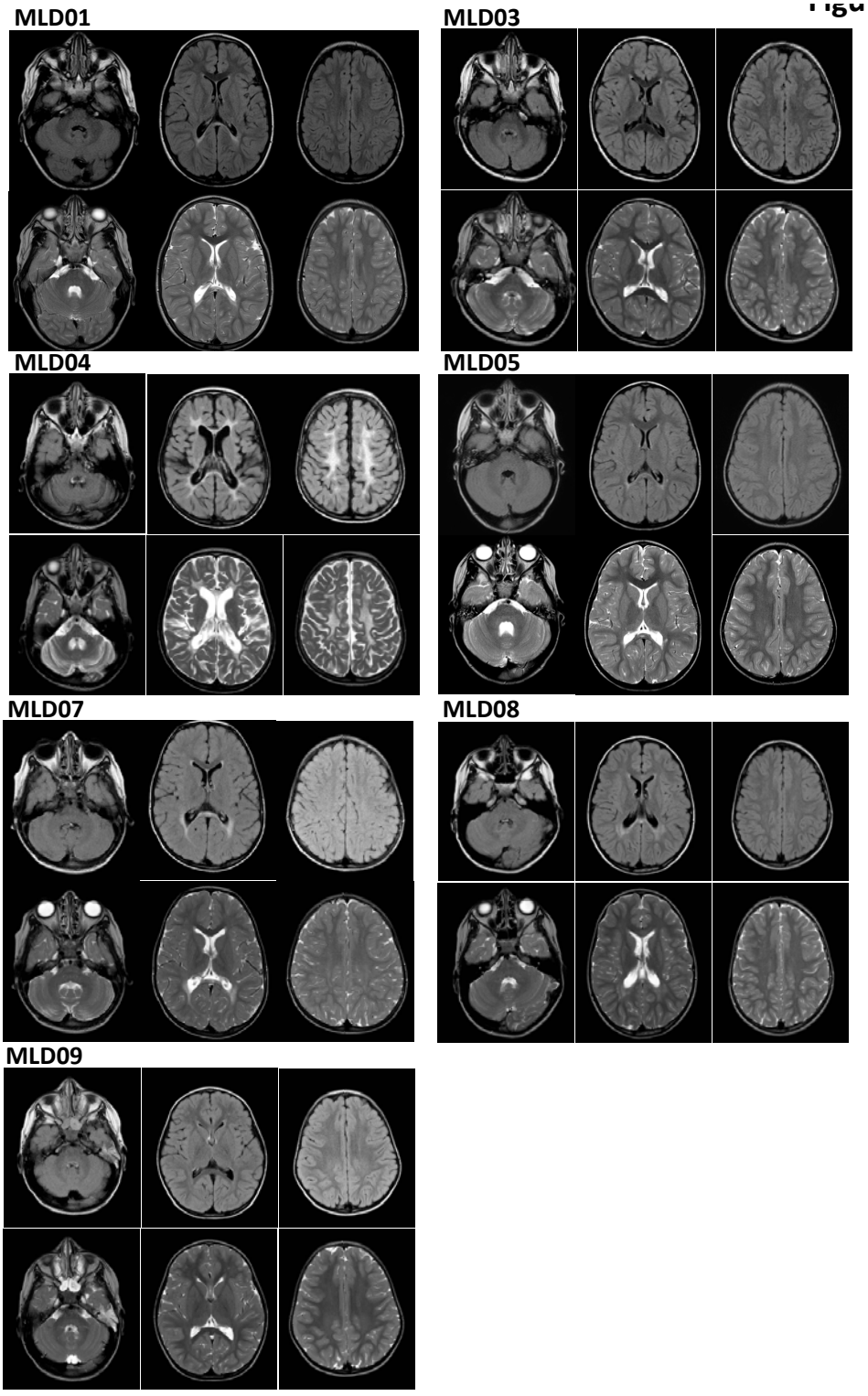


Figure S5. Effect of the treatment on the CNS of early onset patients treated when pre- or early symptomatic. (A) FLAIR MR images (upper row) and Axial T2 weighted fast spin-echo MR images (middle row) obtained from the patients listed in Figure at the last available follow up.

Figure S6

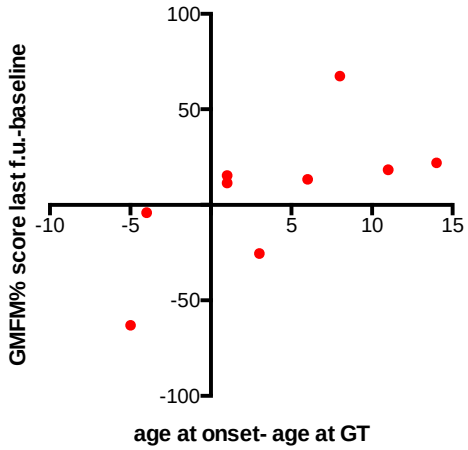


Figure S6. Correlation between the change in GMFM score and the time interval between HSC-GT administration and disease onset.
Correlation between the change in GMFM score from baseline to last follow-up (a positive value indicates that the patient has gained motor skills, a negative value indicates a regression of the patient's motor performance) and the time interval in months between HSC-GT administration and the expected or actual disease onset (a negative value indicates treatment in the presence of symptoms, a positive value indicate presymptomatic treatment; Spearman's r=0·8034; p=0·0138). GMFM=Gross Motor Function Measure. MLD=metachromatic leukodystrophy.
HSC-GT=haemopoietic stem-cell gene therapy.

Figure S7

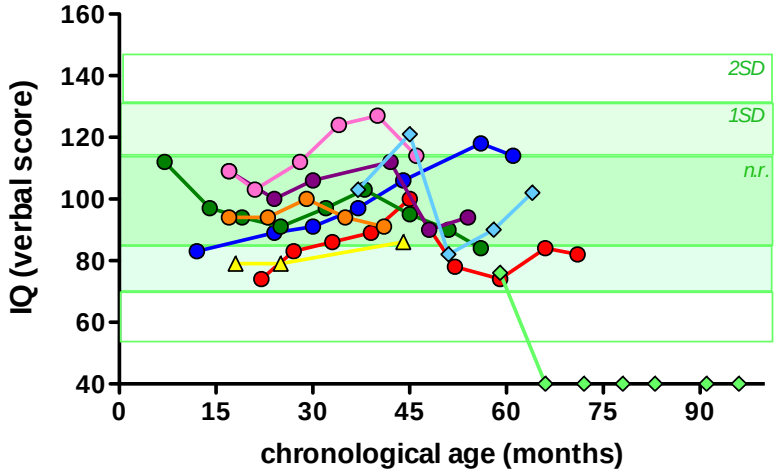


Figure S7. Verbal Intelligence Quotient (IQ) of early onset treated patients. Verbal Intelligence Quotient (IQ), assessed by BSID-III and WPPSI-III according to the age, of early onset pre- or early symptomatic treated patients (see symbol and color-code). The green shaded areas represent the normal range of verbal and performance IQ (up to 2SD).

Table S1

Patient	Mutation	Exon/Intron	cDNA *	Effect of mutation (protein)	Kind of mutation (cDNA)	Kind of mutation (protein)	Classifica
MLD01	ARSA gene mutation 1	Exon 4	c.827C>T	p.Thr276Met	missense substitution	missense	
	ARSA gene mutation 2	Exon 4	c.827C>T	p.Thr276Met	missense substitution	missense	
MLD02	ARSA gene mutation 1	Exon 4	c.736C>T	p.Arg246Cys	missense substitution	missense	
	ARSA gene mutation 2	Exon 4	c.737G>A	p.Arg246His	missense substitution	missense	
MLD03	ARSA gene mutation 1	Exon 2	c.449C>G	p.Pro150Arg	missense substitution	missense	
	ARSA gene mutation 2	Exon 2	c.449C>G	p.Pro150Arg	missense substitution	missense	
MLD04	ARSA gene mutation 1	Exon 8	c.1283C>T	p.Pro428Leu	missense substitution	missense	
	ARSA gene mutation 2	Exon 2	c.383T>G	p.Leu128Arg	missense substitution	missense	
MLD05	ARSA gene mutation 1	Intron 2	c.465+1G>A	r.0	loss of splice donor site	splice donor	
	ARSA gene mutation 2	Intron 5	c.980-1G>A	r.?	loss of splice acceptor site	splice acceptor	
MLD06	ARSA gene mutation 1	Intron 2	c.465+1G>A	r.0	loss of splice donor site	splice donor	
	ARSA gene mutation 2	Intron 4	c.855-1G>A	r.?	loss of splice acceptor site	splice acceptor	
MLD07	ARSA gene mutation 1	Intron 2	c.465+1G>A	r.0	loss of splice donor site	splice donor	
	ARSA gene mutation 2	Intron 2	c.465+1G>A	r.0	loss of splice donor site	splice donor	
MLD08	ARSA gene mutation 1	Exon 8	c.1223_1231delGTGATACCA	p.Ser408_Thr410del	deletion	in-frame deletion	
	ARSA gene mutation 2	Exon 7	c.1150G>A	p.Glu384Lys	missense substitution	missense	
MLD09	ARSA gene mutation 1	Exon 5	c.931G>A	p.Gly311Ser	missense substitution	missense	Uk
	ARSA gene mutation 2	Exon 5	c.931G>A	p.Gly311Ser	missense substitution	missense	Uk

Patient	Mutation	References
MLD01	ARSA gene mutation 1	Harvey et al., Hum Mutat (1993) 2:261-267; Biffi et al., Science (2013) 341, 1233158
	ARSA gene mutation 2	Harvey et al., Hum Mutat (1993) 2:261-267; Biffi et al., Science (2013) 341, 1233158
MLD02	ARSA gene mutation 1	Draghia et al., Hum Mutat (1997) 9:234-242; Gieselmann et al., Hum Mutat (1994) 4:233-242 ; Biffi et al., Science (2013) 341, 1233158
MLD02	ARSA gene mutation 2	Draghia et al., Hum Mutat (1997) 9:234-242; Gieselmann et al., Hum Mutat (1994) 4:233-242 ; Biffi et al., Science (2013) 341, 1233158
MLD03	ARSA gene mutation 1	Cesani et al 2015 (in press)
	ARSA gene mutation 2	Cesani et al 2015 (in press)
MLD04	ARSA gene mutation 1	Polten et al., N Engl J Med (1991) 324:18-22
	ARSA gene mutation 2	Cesani et al 2015 (in press)
MLD05	ARSA gene mutation 1	Polten et al., N Engl J Med (1991) 324:18-22
	ARSA gene mutation 2	Cesani et al 2015 (in press)
MLD06	ARSA gene mutation 1	Polten et al., N Engl J Med (1991) 324:18-22
	ARSA gene mutation 2	Regis et al., Eur J Hum Genet (2004) 12:150-154; Lorioli et al., Gene (2014) 537:348-351
MLD07	ARSA gene mutation 1	Polten et al., N Engl J Med (1991) 324:18-22
MLD07	ARSA gene mutation 2	Polten et al., N Engl J Med (1991) 324:18-22
MLD08	ARSA gene mutation 1	Regis et al., Hum Genet (1998) 102:50-53
MLD08	ARSA gene mutation 2	Barth et al., Hum Mol Genet (1993) 2:2117-2121
MLD09	ARSA gene mutation 1	Kreysing et al., Am J Hum Genet (1993) 53:339-346
MLD09	ARSA gene mutation 2	Kreysing et al., Am J Hum Genet (1993) 53:339-346

Table S1. Patient's molecular characterization.

Exon or Intron involved. * ARSA gene GenBank accession no. NM_000487.5 and NP_000478.3. Mutations are described according to current mutation nomenclature gu (http://www.hgvs.org/mutnomen; [den Dunnen and Antonarakis, 2001]). Effect of mutation (protein)= effect on protein sequence.

Ukn: unknown mutation; 0: severe mutation; R: mutation with residual activity.

Ukn**: for this specific mutation please refer to Appendix S4. MLD09 classification

Mutations were classified as 0 or R as described in Table 3 of Biffi et al 2008 and as Unknown if the mutation was not a splice or null mutation and ARSA activity data from a single allele expression assay were not available