

Supplemental information

Gene therapy supports long-term reconstitution of patient hematopoietic stem cells in deficiency of adenosine deaminase 2

Chiara Rigamonti, Dimitri Bulté, Federica Barzaghi, Cristina Mesa-Nuñez, Luca Basso-Ricci, Emanuela Pettinato, Camilla Visconti, Sara Degl'Innocenti, Francesco Gazzo, Fabrizio Benedicenti, Sergio Arévalo, Pamela Quaranta, Francesca Sanvito, Roberta Caorsi, Francesca Schena, Michela Lupia, Silvia Federici, Antonella Insalaco, Fabrizio De Benedetti, Serena Scala, Maurizio Miano, Francesca Conti, Maria Pia Cicalese, Eugenio Montini, Marco Gattorno, Carlo Dufour, Alessandro Aiuti, and Alessandra Mortellaro

SUPPLEMENTAL MATERIALS

SUPPLEMENTAL METHODS

Vector integration site retrieval and sequencing

Genomic DNA was extracted from human BM and mPB CD34+ cells, and from BM samples harvested from NSGW41 mice. Integration site (IS) retrieval was performed using Sonication Linker-mediated PCR (SLiM-PCR), as previously described^{1,2}. Briefly, ~30 ng of genomic DNA per sample was sonicated and split into three technical replicates. Fragmented DNA was then end-repaired, adenylated, and ligated to a linker cassette using the NEBNext® Ultra™ II DNA Library Prep Kit (New England Biolabs, Cat. n° E7645), following the manufacturer's instruction. Two consecutive PCR amplifications (25 and 10 cycles, respectively) were performed using primers targeting the vector Long Terminal Repeats (LTR) and the linker cassette to selectively amplify vector–genome junctions. Clean-up and concentration steps were included between the two PCR rounds to optimize second amplification.

Primers used in this protocol included sample-specific barcodes and sequencing adapters for multiplexed paired-end sequencing on the MGI platforms. Detailed primer sequences and barcoding information have been previously published¹. A single sequencing library comprising 145 samples was generated and sequenced using the MGI G400 platform, yielding $>1.96 \times 10^8$ raw reads.

Identification of vector integration sites

Vector integration sites (IS) were identified using VISPA2 pipeline³ applied to SLiM-PCR–amplified libraries sequenced with MGI paired-end technology. For each

sequencing library, raw paired-end reads underwent quality control filtering, barcode recognition for sample demultiplexing, and removal of vector sequences. The remaining genomic sequences were aligned to the human reference genome (GRCh37/hg19, February 2019 release).

To quantify the number of genomes corresponding to each clone, we employed the SonicLength approach⁴, which quantifies distinct fragments associated with each IS. This ensured accurate evaluation of clonal abundance within each sample. The final IS dataset included only precisely mapped loci, annotated with the nearest RefSeq gene. Downstream analysis was performed using a new ISAnalytics R package⁵, which integrates VISPA2 output files and supports quality controls, clonal tracking, and inter-sample IS comparison. The full code and documentation are available on GitHub (<https://github.com/calabrialab/ISAnalytics>). To solve “collisions” (identical IS detected across different independent samples), we followed an established method (21), assigning the IS to a single sample based on (i) order of detection or (ii) a ≥ 10 -fold higher abundance in one sample versus the other.

Quality control of the sequencing pools involved: (i) exclusion of samples with raw read counts that were three times lower of the pool’s average, and (ii) removal of suspected cross-sample contamination due to IS collisions, as previously described⁶.

Common insertion site (CIS) analysis was conducted using the Grubbs outlier test⁷ implemented in ISAnalytics. For each patient, the targeting frequency of genes (based on IS within the gene body or ± 100 kb) was normalized to the gene length and log2 transformed. Genes with significantly enriched integration frequency were identified as CIS hits.

Clonal population diversity

An ecological system is maintained stable if the populating species are in equilibrium, suggesting a healthy environment. Analogously, clonal diversity in gene therapy settings can be used as a proxy for hematopoietic health. One commonly used quantitative metric for this is the Shannon diversity index (H-index), which captures both clonal richness (number of distinct clones, or integration sites—IS) and evenness (relative abundance of each clone). H-index is defined as in Equation S1.

This ecological framework has been widely applied to track heterogeneity and complexity of vector-marked cells over time, across tissues, and within differentiated lineages. In this context, ISs are treated as species and their relative abundances reflect clonal prevalence. Sustained high H-index values indicate stable and diverse hematopoiesis, while a sharp decline may signal clonal skewing or potential malignant transformation. In this study, the Shannon diversity index was calculated using the vegan R package and implemented within the ISAnalytics framework to quantify clonal complexity in both in vitro and in vivo settings.

$$H' = - \sum_{i=1}^R p_i \ln p_i$$

Equation S1. Formula for calculating the H-index quantifying clonal complexity, where i denotes a clone (integration site, IS), p_i is the clonal abundance, and R is the total number of clones.

SUPPLEMENTAL FIGURES

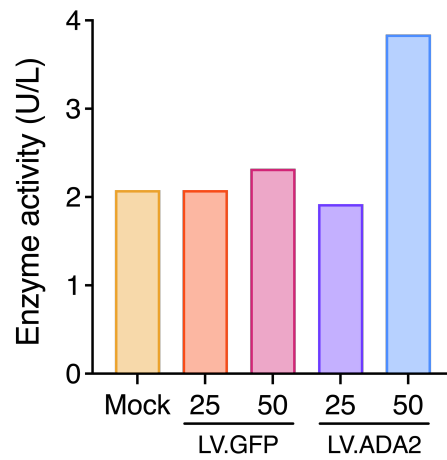


Figure S1. Increased ADA2 activity upon transduction with LV.ADA2 at MOI 50. ADA2 enzymatic activity measured in cell-free supernatants from mPB CD34⁺ cells either mock-transduced or transduced with LV.ADA2 or LV.GFP (control vector) at two multiplicities of infection (MOI 25 and 50).

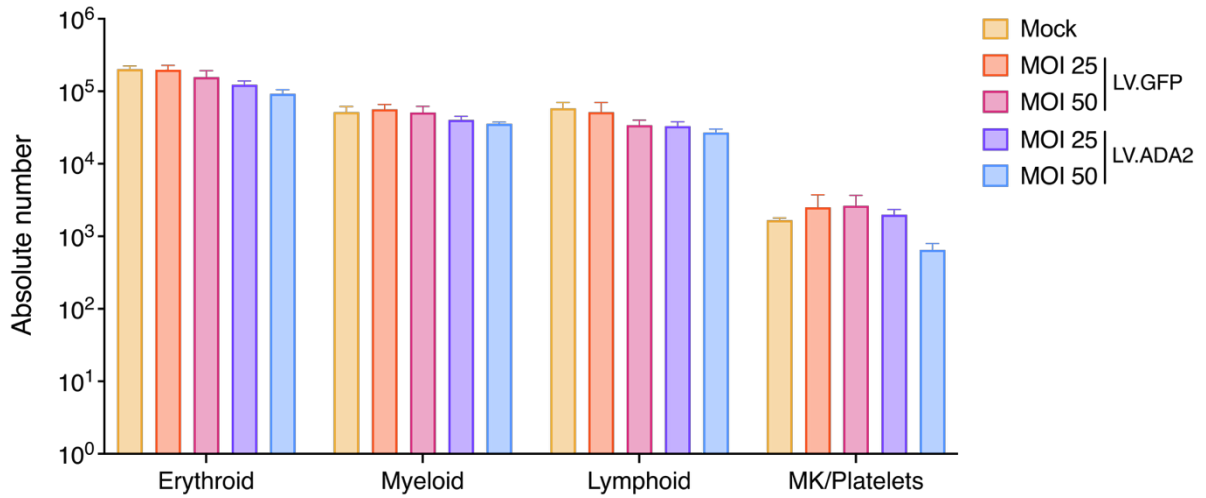


Figure S2. LV-mediated ADA2 overexpression supports normal in vitro differentiation of mPB CD34⁺ cells. In vitro multi-lineage differentiation assay of mPB CD34⁺ cells either mock-transduced or transduced with LV.ADA2 or LV.GFP (control vector) at two multiplicities of infection (MOI 25 and 50). Data are presented as mean±standard error of the mean (SEM).

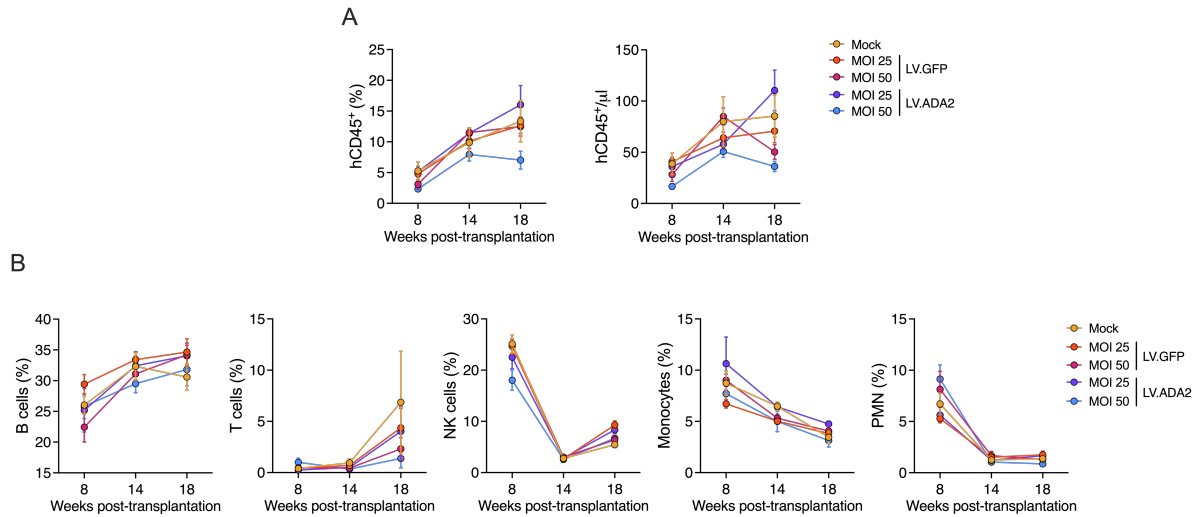


Figure S3. Normal multilineage differentiation of ADA2-transduced mPB CD34⁺ cells from healthy donors in NSGW41 mice. (A) Longitudinal analysis of human chimerism shown as both percentage and absolute number of human CD45⁺ cells in peripheral blood at 8, 14, and 18 weeks post-transplantation (n=5 mice per group). **(B)** Longitudinal analysis of peripheral blood immune subsets—B cells, T cells, NK cells, monocytes, and polymorphonuclear (PMN) cells—at 8, 14, and 18 weeks post-transplantation (n=5 mice per group).

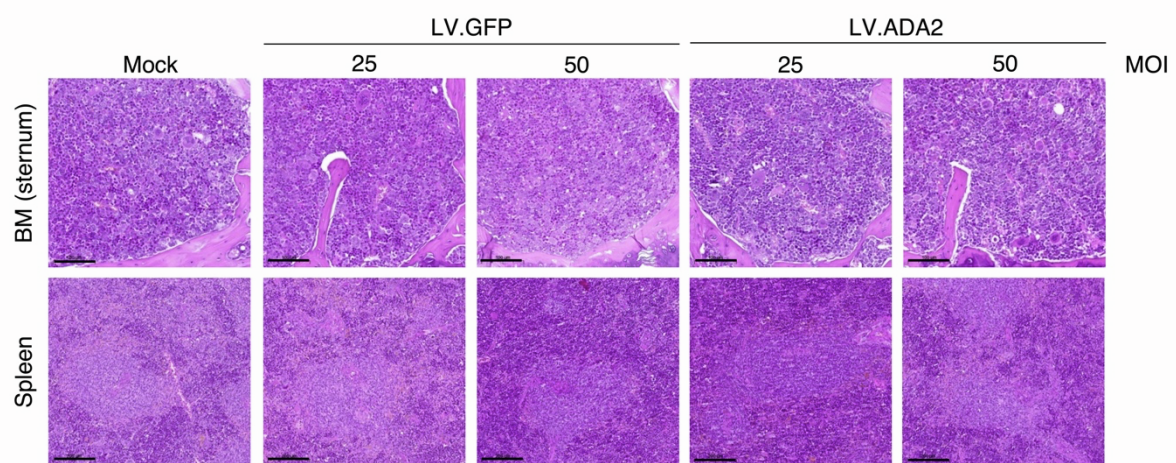


Figure S4. No evidence of in vivo toxicity associated with LV-mediated ADA2 overexpression. Hematoxylin and eosin (H&E) staining of representative sections from the sternum BM and spleen of NSGW41 mice transplanted with mPB CD34⁺ cells from healthy donors (HDs), either mock-transduced or transduced with LV.ADA2 or LV.GFP at a multiplicity of infection (MOI) of 25 or 50. Scale bars: 100 μ m in bone marrow (sternum), 200 μ m in spleen.

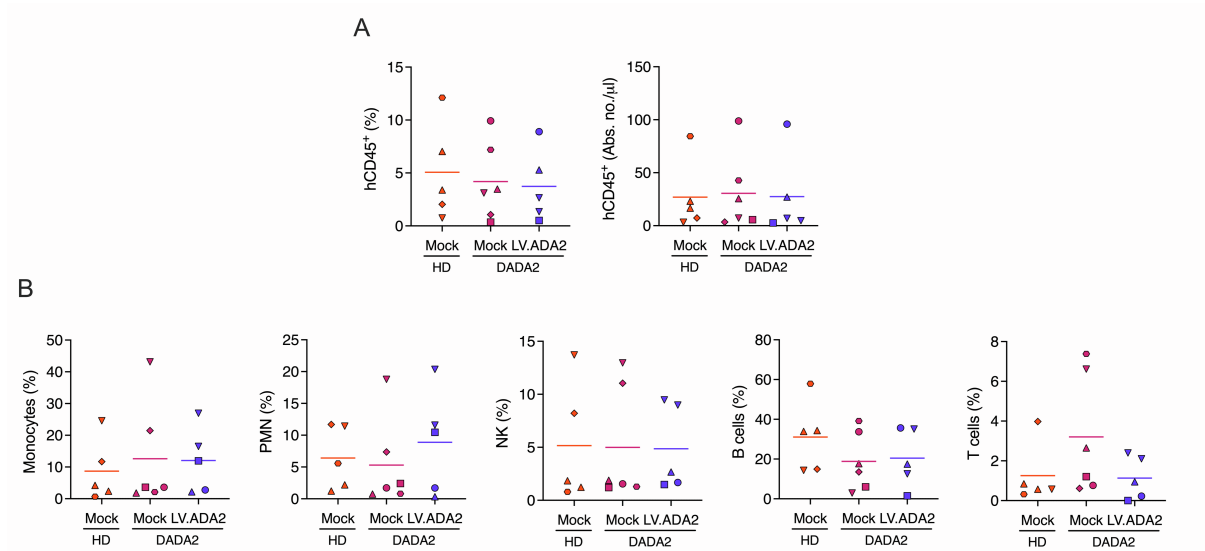


Figure S5. LV.ADA2 gene therapy supports multilineage hematopoietic differentiation of ADA2-transduced patient-derived HSPCs in NSGW41 mice. (A) Human chimerism in peripheral blood was assessed at 20 weeks post-transplantation and is shown as both percentage and absolute number of human CD45⁺ cells per µl. **(B)** Percentages of mature immune cell populations—monocytes, polymorphonuclear cells (PMNs), B cells, T cells, and NK cells—were quantified within the human CD45⁺ compartment of peripheral blood at the same time point. Data include recipients of mock-transduced healthy donor cells, mock-transduced DADA2 cells, and LV.ADA2-transduced DADA2 cells. Bars denote mean.

SUPPLEMENTAL TABLES

Table S1. Number and distribution of unique integration sites in all in vitro and in vivo datasets.

Table S2. Sample characteristics and number of integration sites retrieved

Donor	Cell source	Vector and dose	Tissue	VCN_avg	DNA (ng)	nIS
HD-A	mPB HD CD34+	LV.ADA2 MOI 25	Post-LC	0.40	10.58	346
HD-A	mPB HD CD34+	LV.ADA2 MOI 50	Post-LC	3.40	10.66	2984
HD-A	mPB HD CD34+	LV.GFP MOI 25	Post-LC	0.21	10.66	158
HD-A	mPB HD CD34+	LV.GFP MOI 50	Post-LC	0.89	10.61	712
HD-A	mPB HD CD34+	LV.ADA2 MOI 25	Mouse total BM	0.41	10.45	445
HD-A	mPB HD CD34+	LV.ADA2 MOI 50	Mouse total BM	3.72	10.53	1796
HD-A	mPB HD CD34+	LV.GFP MOI 25	Mouse total BM	0.18	10.52	161
HD-A	mPB HD CD34+	LV.GFP MOI 50	Mouse total BM	0.81	10.45	725
HD1	HD BM CD34+	LV.ADA2 MOI 50	Post-LC	7.00	10.53	2615
HD2	HD BM CD34+	LV.ADA2 MOI 50	Post-LC	15.90	10.46	7661
HD3	HD BM CD34+	LV.ADA2 MOI 50	Post-LC	0.67	10.54	459
HD4	HD BM CD34+	LV.ADA2 MOI 50	Post-LC	0.95	10.37	576
Pt1	DADA2 BM CD34+	LV.ADA2 MOI 50	Post-LC	3.95	10.49	1303
Pt2	DADA2 BM CD34+	LV.ADA2 MOI 50	Post-LC	10.78	10.44	5395
Pt3	DADA2 BM CD34+	LV.ADA2 MOI 50	Post-LC	6.18	10.40	5521
Pt4	DADA2 BM CD34+	LV.ADA2 MOI 50	Post-LC	16.13	10.40	4470
Pt5	DADA2 BM CD34+	LV.ADA2 MOI 50	Post-LC	8.14	10.61	2808
Pt6	DADA2 BM CD34+	LV.ADA2 MOI 50	Post-LC	1.26	10.55	782
HD1	HD BM CD34+	LV.ADA2 MOI 50	Mouse total BM	5.77	10.57	168
HD2	HD BM CD34+	LV.ADA2 MOI 50	Mouse total BM	11.55	10.35	303
HD3	HD BM CD34+	LV.ADA2 MOI 50	Mouse total BM	0.72	10.61	101
HD4	HD BM CD34+	LV.ADA2 MOI 50	Mouse total BM	0.42	10.51	15
Pt1	DADA2 BM CD34+	LV.ADA2 MOI 50	Mouse total BM	3.54	10.40	103
Pt4	DADA2 BM CD34+	LV.ADA2 MOI 50	Mouse total BM	12.80	10.61	47
Pt5	DADA2 BM CD34+	LV.ADA2 MOI 50	Mouse total BM	2.01	10.53	133
Pt6	DADA2 BM CD34+	LV.ADA2 MOI 50	Mouse total BM	0.50	10.40	105

Healthy donors (HD) or DADA2 patients (Pt) (Donor column) were used as source of mobilized peripheral blood (mPB) or bone marrow (BM) derived CD34+ cells (column cell source) and transduced with LV.ADA2 or LV.GFP at a MOI of 25 or 50 (column Vector and Dose). Integrations were retrieved from cells after 14-day in vitro liquid culture (Post-LC) or BM cells from NSGW41 recipient mice at 20 weeks post-transplantation (column Tissue). The table reports the average vector copy number (VCN), the amount of genomic DNA used for SLiM-PCR, and the number of unique integration sites (nIS) identified per condition are indicated.

References

1. Cesana, D., Calabria, A., Rudilosso, L., Gallina, P., Benedicenti, F., Spinozzi, G., Schioli, G., Magnani, A., Acquati, S., Fumagalli, F., et al. (2021). Retrieval of vector integration sites from cell-free DNA. *Nat Med* 27, 1458-1470. 10.1038/s41591-021-01389-4.
2. Firouzi, S., Lopez, Y., Suzuki, Y., Nakai, K., Sugano, S., Yamochi, T., and Watanabe, T. (2014). Development and validation of a new high-throughput method to investigate the clonality of HTLV-1-infected cells based on provirus integration sites. *Genome Med* 6, 46. 10.1186/gm568.
3. Spinozzi, G., Calabria, A., Brasca, S., Beretta, S., Merelli, I., Milanesi, L., and Montini, E. (2017). VISPA2: a scalable pipeline for high-throughput identification and annotation of vector integration sites. *BMC Bioinformatics* 18, 520. 10.1186/s12859-017-1937-9.
4. Berry, C.C., Gillet, N.A., Melamed, A., Gormley, N., Bangham, C.R., and Bushman, F.D. (2012). Estimating abundances of retroviral insertion sites from DNA fragment length data. *Bioinformatics* 28, 755-762. 10.1093/bioinformatics/bts004.
5. Pais, G., Spinozzi, G., Cesana, D., Benedicenti, F., Albertini, A., Bernardo, M.E., Gentner, B., Montini, E., and Calabria, A. (2023). ISAnalytics enables longitudinal and high-throughput clonal tracking studies in hematopoietic stem cell gene therapy applications. *Brief Bioinform* 24. 10.1093/bib/bbac551.
6. Biffi, A., Montini, E., Lorioli, L., Cesani, M., Fumagalli, F., Plati, T., Baldoli, C., Martino, S., Calabria, A., Canale, S., et al. (2013). Lentiviral hematopoietic stem cell gene therapy benefits metachromatic leukodystrophy. *Science* 341, 1233158. 10.1126/science.1233158.

7. Biffi, A., Bartolomae, C.C., Cesana, D., Cartier, N., Aubourg, P., Ranzani, M., Cesani, M., Benedicenti, F., Plati, T., Rubagotti, E., et al. (2011). Lentiviral vector common integration sites in preclinical models and a clinical trial reflect a benign integration bias and not oncogenic selection. *Blood* 117, 5332-5339. 10.1182/blood-2010-09-306761.