YMTHE, Volume *25*

Supplemental Information

Lentiviral Transfer of γ-Globin with Fusion Gene NUP98-HOXA10HD Expands Hematopoietic Stem Cells and Ameliorates Murine β-Thalassemia

Hui Fen Zhao, Allistair Abraham, Yoon-Sang Kim, Yong-Dong Wang, Tamara Pestina, Jun Zhan, Keith Humphries, Arthur W. Nienhuis, and Derek A. Persons

Figure S1: **Construction of lentiviral vectors**

(a) Proviral map of the integrated lentiviral γ- globin-MNDU3-NA10HD vector which contains the composite 3.1 kb of transcriptional regulatory sequences from the β-globin LCR as indicated, from HS4, HS3, and HS2, coupled with a 130 bp β-globin promoter and genomic γ-globin sequences. As indicated, NA10HD expression (γ-globin/NA10HD) is driven by the MND U3 promoter.

(b) Proviral map of the control vector where the vector in Figure S1 in which NA10HD was replaced by mCherry.

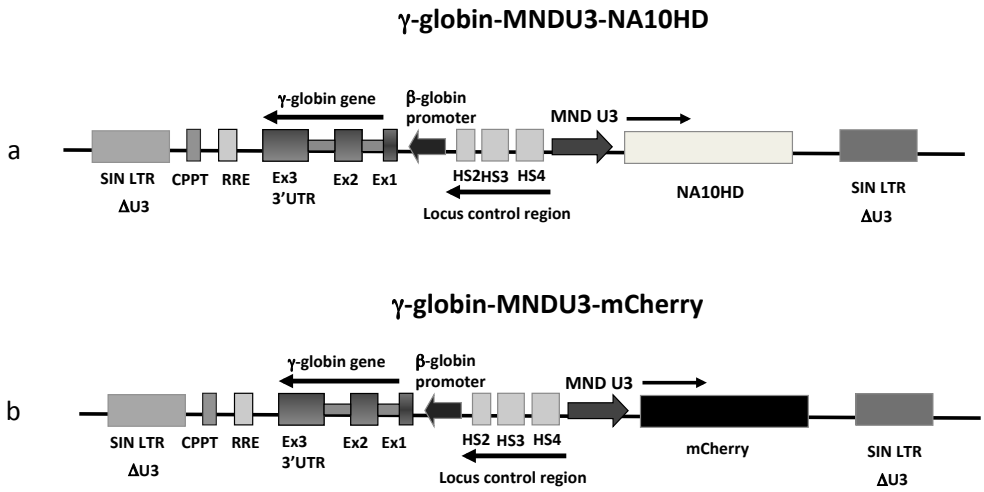

γ-globin-MNDU3-NA10HD
a
γ-globin gene
β-globin promoter
MND U3
SIN LTR ΔU3
CPPT
RRE
Ex3 3'UTR
Ex2
Ex1
HS2
HS3
HS4
Locus control region
NA10HD
SIN LTR ΔU3
γ-globin-MNDU3-mCherry
b
γ-globin gene
β-globin promoter
MND U3
SIN LTR ΔU3
CPPT
RRE
Ex3 3'UTR
Ex2
Ex1
HS2
HS3
HS4
Locus control region
mCherry
SIN LTR ΔU3

Figure S1

Figure S2: **NA10HD protein expression in 293T cells transduced with the γ-globin /NA10HD vector**

(a) IgG1 isotype antibody only as control, no FLAG positive cells were detected.

(b) When the anti FLAG antibody was used to detect NA10HD positive cells, 14.7% shifted into the positive gate showing expression of the NA10HD-FLAG.

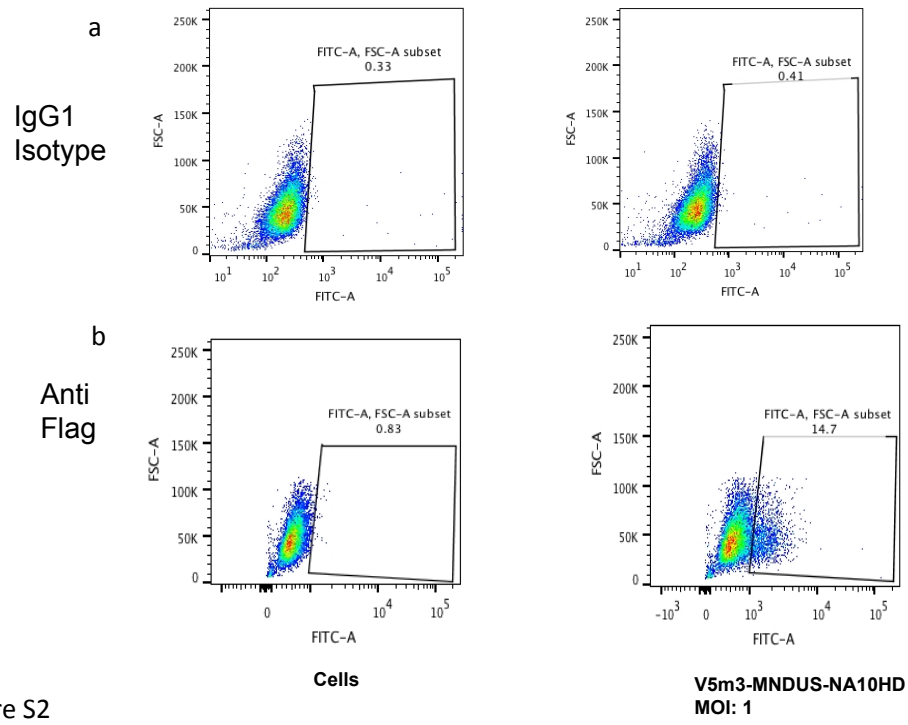

a
IgG1
Isotype
b
Anti
Flag
FITC-A, FSC-A subset
0.33
FITC-A, FSC-A subset
0.41
FITC-A, FSC-A subset
0.83
FITC-A, FSC-A subset
14.7
FSC-A
FITC-A
Cells
V5m3-MNDUS-NA10HD
MOI: 1

Figure S2

Figure S3: **Analysis of genomic DNA from cells transduced with the γ-globin/NA10HD and γ-globin/mCherry vectors**

(a) Genomic DNA digested by using EcoRI restriction enzyme was derived from spleen colonies from β-thalassemic BM cells transduced with the γ-globin/NA10HD vector. (b) Genomic DNA digested by using EcoRI restriction enzyme was derived from spleen colonies from β-thalassemic BM cells transduced with the γ-globin/mCherry vector. The number of vector insertion sites for each spleen colony is indicated at bottom of the gel.

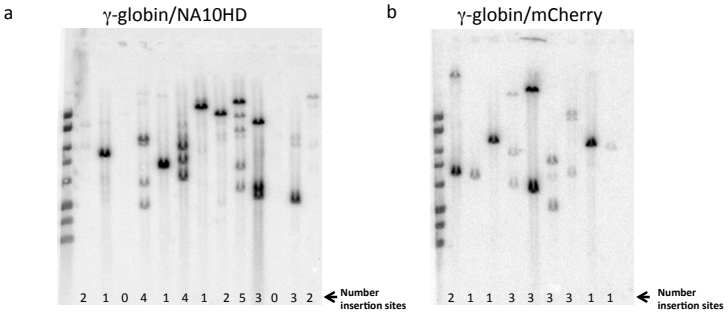

a
γ-globin/NA10HD
2 1 0 4 1 4 1 2 5 3 0 3 2
Number insertion sites
b
γ-globin/mCherry
2 1 1 3 3 3 3 1 1
Number insertion sites

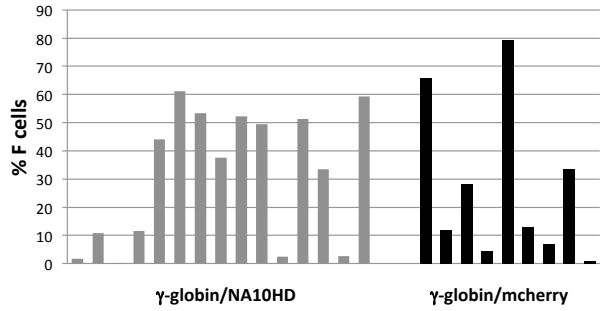
Figure S3

Figure S4: **F positive cells were detected in spleen colony forming cells in both γ-globin /NA10HD and γ-globin/mCherry mice**

(a) Percentages of F cells obtained from individual spleen colonies were quantified and compared between the γ-globin/NA10HD and the γ-globin/mCherry groups.

(b) Both HbF and mCherry positive cells were quantified and compared to the γ-globin/mCherry control group.

a


90
80
70
60
50
40
30
20
10
0
% F cells
γ-globin/NA10HD
γ-globin/mcherry

b

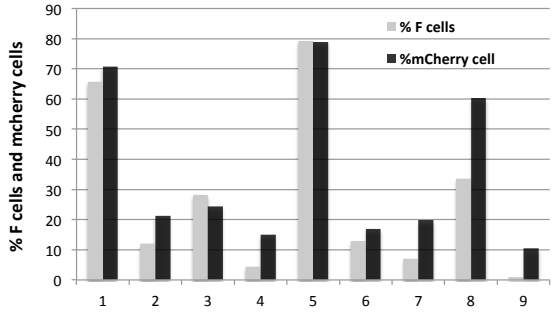

90
80
70
60
50
40
30
20
10
0
% F cells and mcherry cells
% F cells
%mCherry cell
1
2
3
4
5
6
7
8
9

Figure S4

Figure S5: **Increase in HbF and hemoglobin levels three and six months post-transplantation**

(a and b) HbF protein levels at three and six months following transplant were quantified by using HPLC in both the γ-globin/NA10HD (a) and the γ-globin/mCherry (b) groups. (c and d) The individual animal hemoglobin (HB) levels at three and six months after transplantation were quantified CBC in either γ-globin/NA10HD groups (c) or γ-globin/mCherry (d) groups.

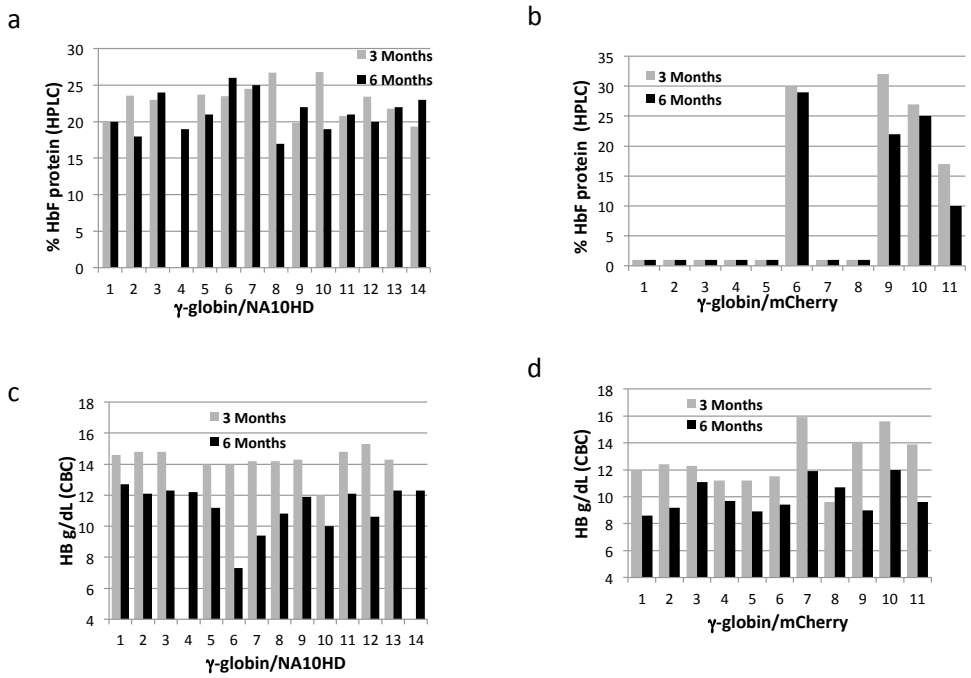

a
% HbF protein (HPLC)
3 Months
6 Months
γ-globin/NA10HD
b
% HbF protein (HPLC)
3 Months
6 Months
γ-globin/mCherry
c
HB g/dL (CBC)
3 Months
6 Months
γ-globin/NA10HD
d
HB g/dL (CBC)
3 Months
6 Months
γ-globin/mCherry

Figure S5

Figure S6: **Comparison of HbF protein and hemoglobin levels in two experiments using γ-globin/NA10HD**

(a) Comparison of the HbF protein levels at 6 months in the γ-globin/NA10HD group and 6 months in the γ-globin/NA10HD 2 groups.

(b) Comparison of the HB levels at 6 months in the γ-globin/NA10HD and at 6 months in the γ-globin/NA10HD2 groups.

(c) RBC lysates derived from individual transplanted animals run on a cellulose acetate gel. The first two and last two samples were either BL6 or β-thalassemia mice as shown. The rest of samples were γ-globin/NA10HD transplanted mice.

(d) Correlation between HB levels and RBCs in the γ-globin/NA10HD2 group.

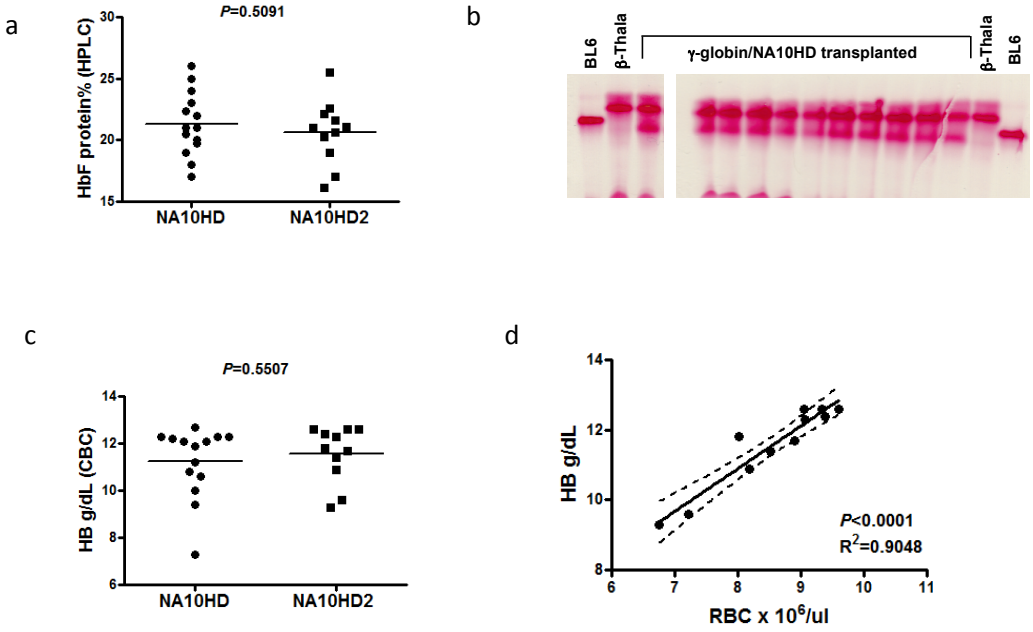

a
P=0.5091
HbF protein% (HPLC)
30
25
20
15
NA10HD
NA10HD2
b
BL6
β-Thala
γ-globin/NA10HD transplanted
β-Thala
BL6
c
P=0.5507
HB g/dL (CBC)
14
12
10
8
6
NA10HD
NA10HD2
d
HB g/dL
14
12
10
8
6
7
8
9
10
11
P<0.0001
R2=0.9048
RBC x 10^6/ul

Figure S6

Figure S7: **Comparison of the total WBC count in three individual experiments using β-thalassemia mice that either received lentiviral-transduced bone marrow or were not transplanted**

(a) Total WBC count in the β-thalassemia mice (Thala) compared with these transuced with the γ-globin vector (V5m3-Thala) or the γ-globin/NA10HD (NA10HD) vector transplanted β-thalassemia mice, as well as γ-globin vector transplanted BL6 mice (V5m3-BL6).

(b) Total WBC count in β-thalassemia mice (Thala) was compared in animals transplanted with un-transduced "Mock" β-thalassemic marrow (V5m3-Thala, V5m3-BL6, and NA10HD).

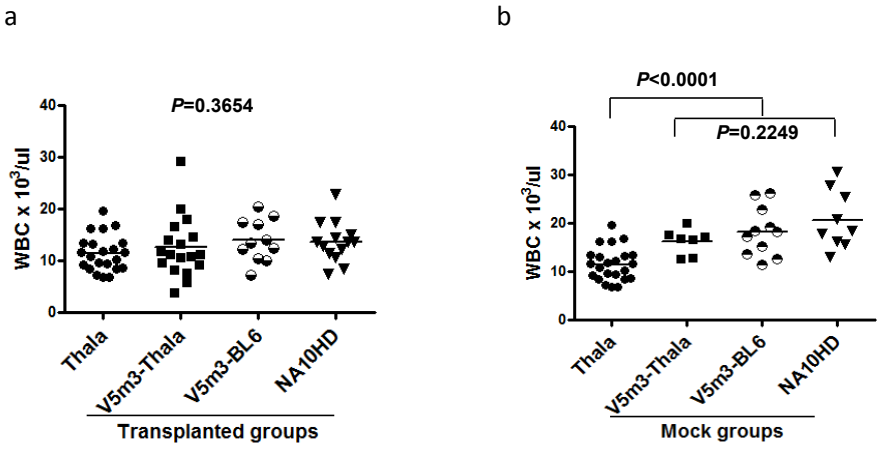

a
b
P=0.3654
P<0.0001
P=0.2249
WBC x 10^3/ul
Thala
V5m3-Thala
V5m3-BL6
NA10HD
Transplanted groups
Mock groups

Figure S7

Figure S8: **HbF protein levels in the γ-globin/NA10HD group were preserved in the secondary transplant groups**

The HbF levels in RBC cells derived from one secondary transplant recipient were separated on a cellulose acetate gel. The first lane on the gel is the Bl6 wild type control.

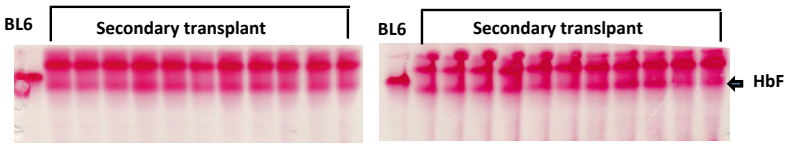


Figure S8

Figure S9: **Relative frequency of sequencing reads of integration sites in mouse ID 4862**

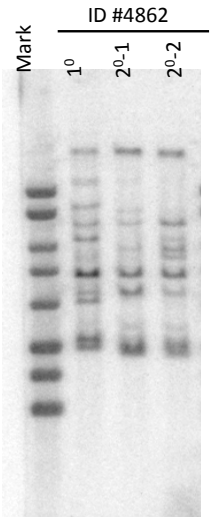
The number of total counted reads for ID4862 was approximately 100,000. There were a total of 82 unique integration sites detected by merging results from three experiments that used different restriction enzymes. Among this collection of restriction sites, there were 15 sites with frequency greater than 1% of the total reads, and 12 of them (80%) were located in RefSeq genes. While most of these 12 sites were in intronic regions, the coding sequences of two genes, 2610507B11Rik in 4% and Capn15 1% of frequency, were interrupted; however there is no indication or previous report that either of these genes playing a role in leukemia. For those genes with an integration site in an intronic region, Eng had an integration site frequency of 1.73%. It was listed as an oncogene in the Uniport database. The relative location of the integration site was further detailed in the second intron of the gene and it was almost 2kb away from the nearest splice site, a region that is not usually considered associated with splicing junction sequences. We concluded that it was unlikely for this integration to have disrupted the protein structure or to cause leukemia as a consequence. The virus insertion sites in the secondary transplanted animals did not display any further sites of amplification when compared to the primary transplanted mouse (Figure S8 A).

(a) 1^0 represents primary transplanted animals that received γ-globin/ NA10HD vector transduced cells 6 months previously and the bone marrow genomic DNA was digested by using EcoR1 restriction enzyme for the Southern blot analysis. 2^0-1 and 2^0-2 were from the secondary transplanted animals that received bone marrow cells from a1^0

primary animal. The genomic DNA used for Southern blot analysis from 2^0 secondary animals was obtained three months after transplantation.

(b) Relative frequency of sequencing reads of integration sites in mouse #4862. The relative frequencies of sequencing reads corresponding to unique integration sites were summarized according to the normalized read counts of experiments using three different restriction enzymes. Fifteen of a total of 82 integration sites, with frequency greater than 1%, were shown with their associated RefSeq genes indicated (*, gene located within 5kb; **, within 10kb; and n/a, no RefSeq gene within 50kb). Sixty-seven integration sites represented by less than 1% of the total reads were pooled and shown in the top of the bar.

a


ID #4862
Mark
1°
2°-1
2°-2

b

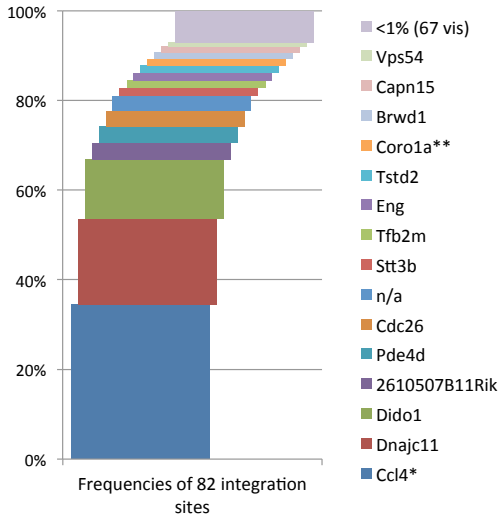

100%
80%
60%
40%
20%
0%
Frequencies of 82 integration sites
<1% (67 vis)
Vps54
Capn15
Brwd1
Coro1a**
Tstd2
Eng
Tfb2m
Stt3b
n/a
Cdc26
Pde4d
2610507B11Rik
Dido1
Dnajc11
Ccl4*

Figure S9

Figure S10al: **Classification of viral vector γ-globin/NA10HD integration sites**

Using the mouse RefSeq genes (GRCm38/mm10, UCSC), the 715 integration sites were classified into seven different categories and the percentages of total sites for each category is shown.

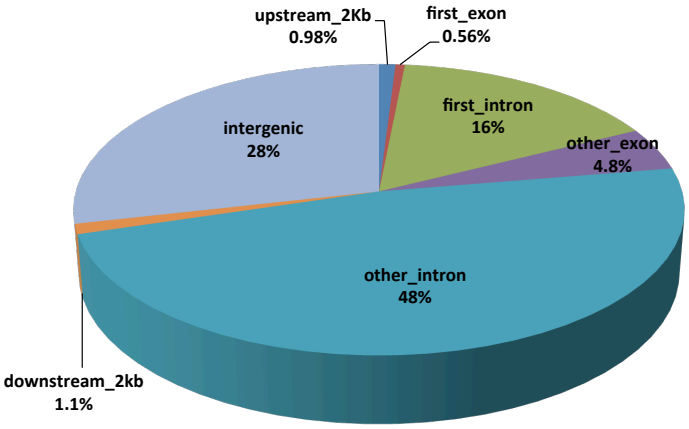

upstream_2Kb
0.98%
first_exon
0.56%
first_intron
16%
other_exon
4.8%
intergenic
28%
other_intron
48%
downstream_2kb
1.1%

Figure S10