OMTM, Volume *17*

Supplemental Information

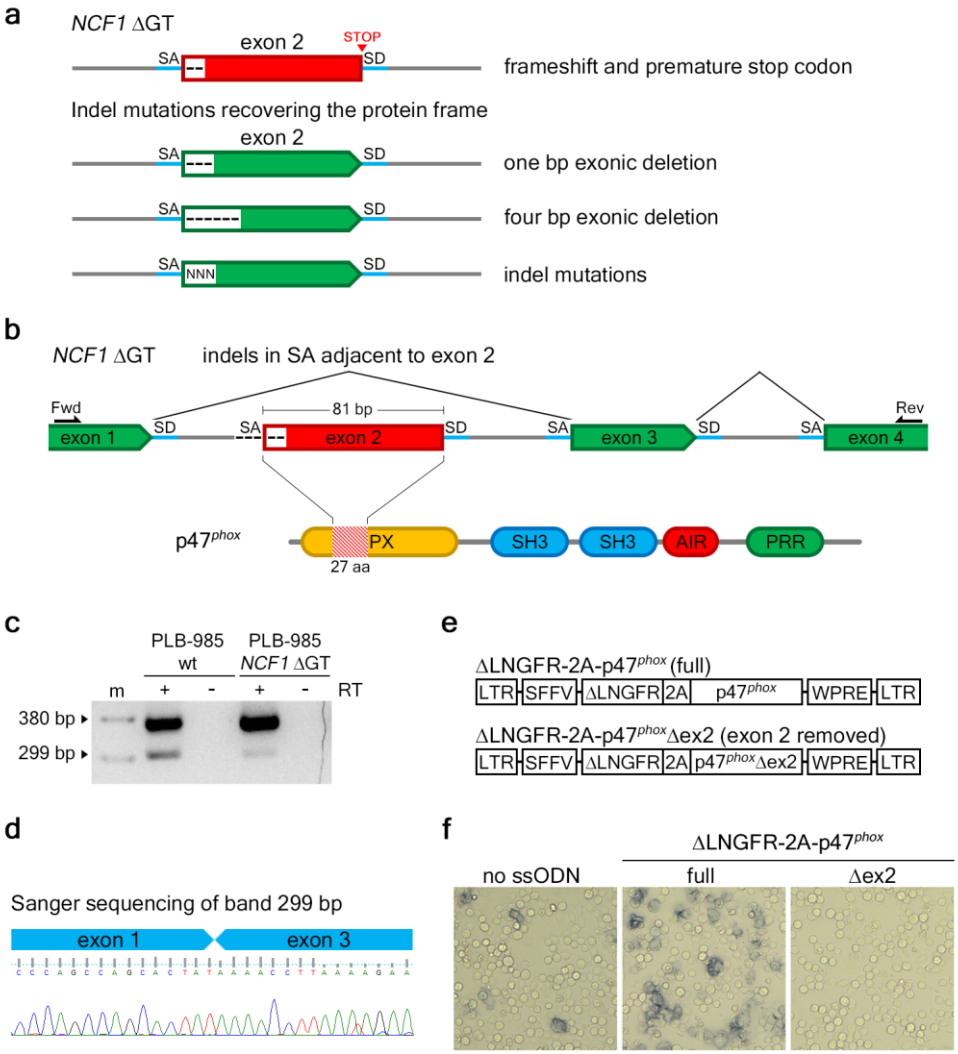
CRISPR-Directed Therapeutic Correction at the *NCF1* Locus Is Challenged by Frequent Incidence of Chromosomal Deletions

Dominik Wrona, Oleksandr Pastukhov, Robert S. Pritchard, Federica Raimondi, Joëlle Tchinda, Martin Jinek, Ulrich Siler, and Janine Reichenbach

Supplemental Figures

Non-homologous end joining (NHEJ)-based repair of CRIPSR-Cas9-induced double-strand breaks (DSBs) at the ΔGT mutation within mutated *NCF1* locus and its pseudogenes generates a variety of indel mutations. Some of these mutations lead to recovery of the *NCF1* gene open reading frame, and may result in expression of a truncated but functional version of p47^{phox} (see **Table S4, and Figure S1**).

Figure S1



Reconstitution of NADPH oxidase activity upon CRISPR-Cas9-treatment of PLB-985 *NCF1* ΔGT cell line in the absence of corrective template

(**a**) Scheme of exon 2 (red) of the *NCF1* locus, carrying the ΔGT mutation (--), with indicated location of stop codon, as well as of *NCF1* exon 2 with possible open reading frame-recovering indel mutations: --- (deletion of 1bp truncates p47^{phox} by 1 amino acid) or ------ (deletion of 4 bp truncates p47^{phox} by 2 amino acids), NNN (insertion of 2 bp, or complex indel mutation).

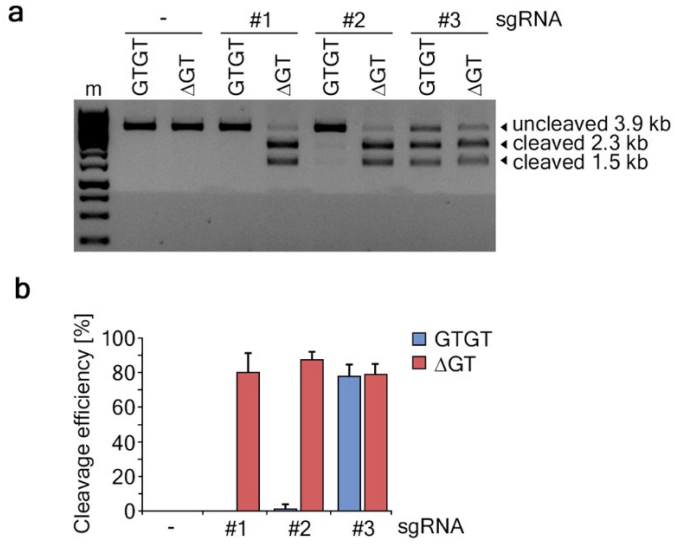
(**b**) Scheme of *NCF1* gene fragment including exons 1-4 (top), as well as scheme of p47^{phox} protein with distinct protein domains. *NCF1* exon 2 consists of 81 bp that encode for 27 amino acids of the p47^{phox} membrane phospholipid binding phox homology (PX) domain. Since the ΔGT mutation is localized at the beginning of exon 2, CRISPR-Cas9-induced indel mutations can potentially lead to disruption of the adjacent splicing acceptor site (SA), and as a consequence to alternative splicing of *NCF1* transcript lacking exon 2 (exon skipping). This in turn may lead to expression of a truncated, yet in-frame version of p47^{phox}. Other depicted domains of p47^{phox} are Src homology 3 domain (SH3), auto-inhibitory region (AIR), and proline-rich region (PRR).

(**c**) Alternative splicing of *NCF1* mRNA in granulocytes differentiated from the PLB-985 *NCF1* ΔGT cell line leads to the presence of *NCF1* transcript without exon 2. PLB-985 wt and PLB-985 *NCF1* ΔGT were differentiated into

granulocytes, and total RNA was isolated from the cells (RNeasy Mini Kit, Qiagen). The presence of correctly spliced *NCF1* transcripts (380bp RT-PCR product), and of exon 2 deficient transcripts (299bp RT-PCR product) in PLB-985 wt and PLB-985 *NCF1* ΔGT was confirmed by reverse transcription. PCR reaction was performed using primers indicated in b: forward (Fwd: ACA CCT TCA TCC GTC ACA TCG) and reverse (Rev: CGT GGG GAG CTT GAG GTC AT, following cDNA generation with RevertAid RT Reverse Transcription Kit (Thermo Fisher Scientific).
(**d**) Sanger sequencing of RT-PCR products confirmed the presence of exon 2 deficient transcripts. (**e**) γ-retroviral vectors used to express full, or exon 2 deficient p47^{phox} in PLB-985 *NCF1* ΔGT cells. The expression cassette between long terminal repeats (LTR) was composed of spleen focus-forming virus (SFFV) promoter, a truncated low-affinity nerve growth factor receptor (ΔLNGFR) surface protein marker, allowing for isolation of successfully transduced cells, and full, or exon 2 deficient *NCF1* cDNA, fused to ΔLNGFR by a 2A self-cleaving peptide.[45] WPRE, woodchuck hepatitis virus posttranscriptional regulatory element.
(**f**, middle and right images) PLB-985 *NCF1* ΔGT cells were transduced with γ-retroviral vectors (schemes in **e**), followed by ΔLNGFR-expressing cells enrichment by fluorescence activated cell sorting (FACS) 24 hours post-transduction, and cultured. The NADPH oxidase activity in differentiated cells was analyzed by NBT test.
(**f**, left image) NBT assay of nucleofected PLB-985 *NCF1* ΔGT cells. Cells were nucleofected as described in Methods of the main text, but without the corrective ssODN template (no ssODN). NBT test was performed on the differentiated bulk culture, and demonstrated reconstitution of ROS production in individual cells.

The results shown in **f** indicate that use of a corrective template for restoration of the NADPH oxidase function is not indispensable (**f**, left image), as suggested by the results observed in another study on X-CGD.[20] On the other hand, *NCF1* exon 2 deficient transcript does not recover the function of the NADPH oxidase (**f**, right image). Only the cells transduced with a vector expressing complete p47^{phox} (**f**, full) reconstituted the activity of the enzyme, whereas the truncated form of p47^{phox} that had exon 2 removed (**f**, Δex2) did not rescue the CGD phenotype. These results indicate that the PX domain, encoded partially by *NCF1* exon 2 is elementary for p47^{phox} protein function, and that expression of p47^{phox} lacking exon 2 due to disruption of the splicing site adjacent to the ΔGT mutation by gene-editing-induced indel mutations does not contribute to the reconstitution of NADPH oxidase function. Therefore, reconstitution of NADPH oxidase function upon CRISPR-Cas9 treatment without corrective template must be the result of created indel mutations, which lead to the expression of an in-frame modified version of p47^{phox}.

Figure S2



Plasmid DNA cleavage assay to test specificity of sgRNAs

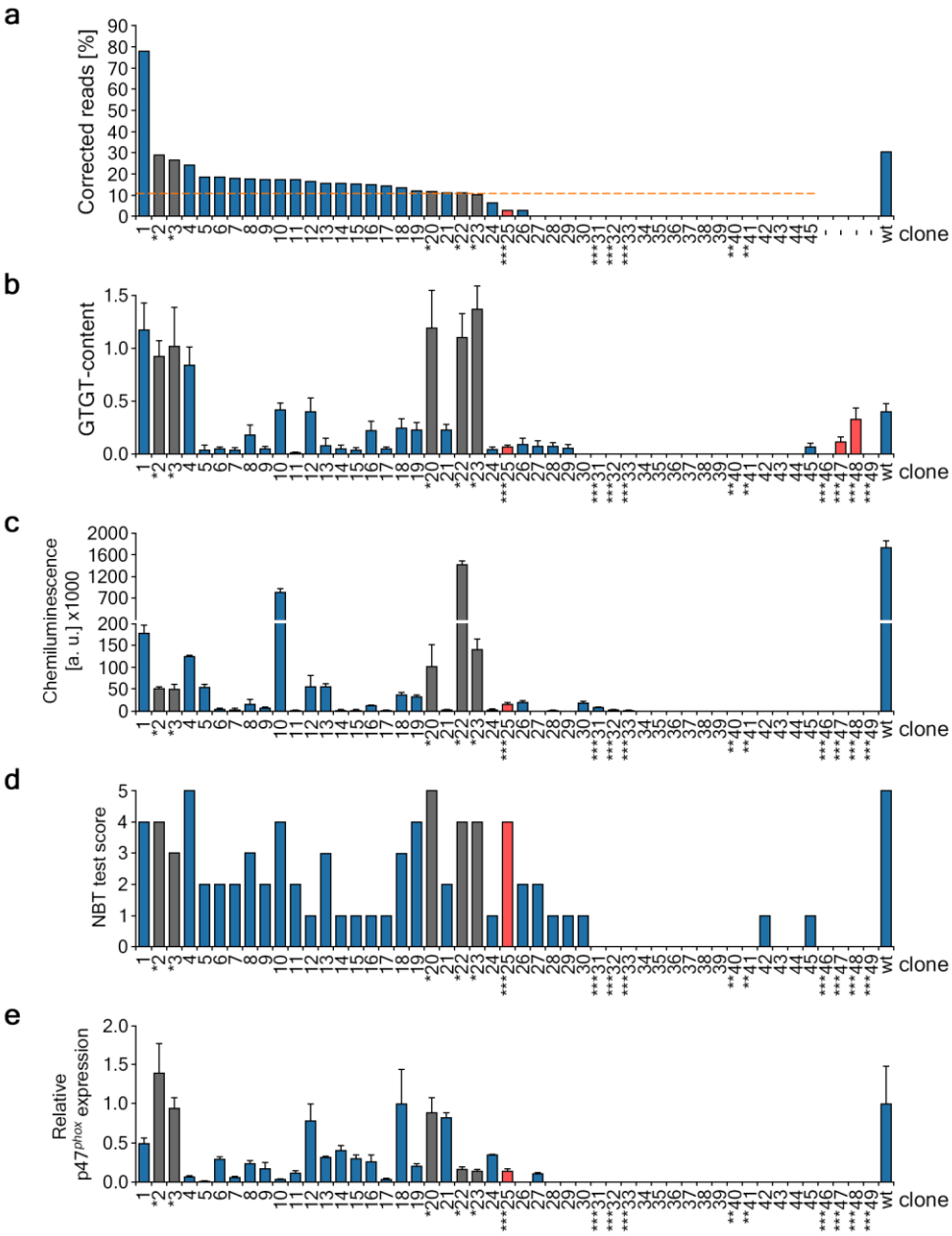
In order to test the specificity of sgRNAs toward the sequence of *NCF1* carrying the GTGT-tetranucleotide or ΔGT mutation, two plasmids with the fragment of the *NCF1* gene or pseudogene sequence spanning the GTGT-tetranucleotide or ΔGT mutation, respectively, were subjected to the plasmid DNA cleavage assay.[46]

Briefly, the plasmids were linearized by restriction digestion with SmaI restriction enzyme (New England Biolabs, Frankfurt am Main, Germany) and subsequently incubated with a 1.5 µM ribonucleoprotein complex of purified SpCas9 and *in vitro* T7 polymerase-transcribed sgRNA. The plasmid cleavage reaction was performed at 37°C for 5 hours in the cleavage buffer (20 mM HEPES pH 7.5, 100 mM KCl, 2 mM $MgCl_2$, 5% (v/v) glycerol, 1 mM DTT, 0.5 mM EDTA). The reaction was stopped with 40 mM EDTA, pH 8.0 followed by sgRNA and Cas9 protein removal with RNase A (Thermo Fisher Scientific, Reinach, Switzerland) and Proteinase K (Roche Diagnostics AG, Rotkreuz, Switzerland) treatment, respectively. RNase A was added to the blocked reaction to reach the final concentration of 0.8 mg/mL and incubated for 1 hour at room temperature. Subsequently, Proteinase K was added to the sample at the final concentration of 1.4 mg/mL and incubated for 30 minutes at room temperature. The samples were developed by agarose electrophoresis.

(**a**) Results of the plasmid DNA cleavage assay developed by agarose electrophoresis. GTGT and ΔGT indicate linearized plasmids (uncleaved 3.9 kb) carrying the GTGT-tetranucleotide or ΔGT mutation, respectively. The bands 'cleaved 2.3 kb' as well as 'cleaved 1.5 kb' result from sgRNA-mediated Cas9 digestion.

(**b**) Quantification of the cleavage efficiency based on band intensities using ImageJ.[43] Bars represent mean cleavage efficiency plus standard deviation of three independent sgRNA transcriptions, ribonucleoprotein complex formations and digestion reactions.

Figure S3



Genetic and functional characterization of individual clones of CRISPR-Cas9-treated PLB-985 *NCF1* ΔGT cell line

Color legend: bars of clones that exhibited no chromosomal aberrations (blue, no asterisk); clones with identified 0.5 kb microdeletion at the CRISPR-Cas9 on-target site (see **Figure 2A**; grey, one asterisk); clones with identified chromosomal deletions between *NCF1* loci (red, two asterisks: deletion identified between *NCF1* and *NCF1C;* three asterisks: deletion identified between *NCF1B* and *NCF1C*).

(**a**) Percentage of SMRT sequencing reads with corrected ΔGT mutation. Mean percentage of corrected reads in Cas9-expressing clones (horizontal orange dashed line).

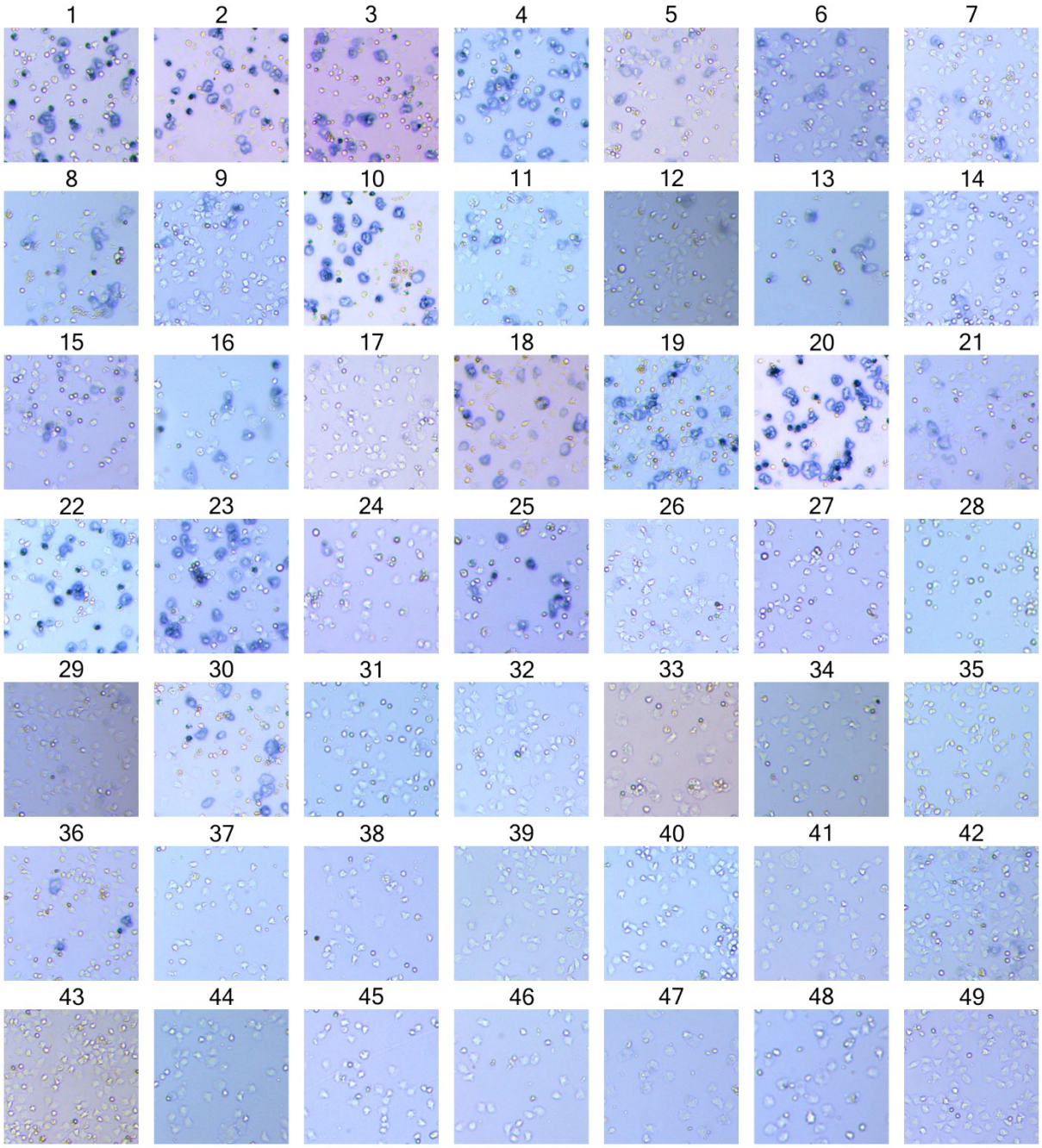
(**b**) PCR-RFLP analysis of PCR co-amplification products with reconstituted BsrG1 restriction site. The GTGT-content determines how many *NCF1* loci were corrected in the pool of *NCF1* loci present in the genome of CRISPR-Cas9-treated clones. Nota bene, the presence of 0.5 kb microdeletion caused artificially high GTGT-values, as the deletion removes the reverse primer binding site utilized in PCR-PFLP analysis.

(**c**) Chemiluminescence assay performed on clones after differentiation to granulocytes. Bars: area under the curve within the first 30 minutes after activation with PMA, representing the activity of the NADPH oxidase.

(**d**) NBT test performed on clones after differentiation to granulocytes. The NADPH oxidase-dependent deposition of formazan within CRIPSR-Cas9-treated cells was assessed microscopically (NBT test score 0-5; 0: no activity, 5: highest activity; see **Figure S4**).
(**e**) Relative expression of $p47^{phox}$ protein by western blot in corrected clones differentiated to granulocytes, as compared to PLB-985 wt cells. Expression of the phagocytic NADPH oxidase protein components ($gp91^{phox}$, $p22^{phox}$, $p67^{phox}$, $p47^{phox}$, $p40^{phox}$) in PLB-985 cell line is induced by granulocytic differentiation.[47] Band intensities of $p47^{phox}$ were controlled for differentiation status by normalization of the signal by signal of $p67^{phox}$ (see representative western blot membrane in **Figure S5**).
Data represented as means of three measurements plus standard deviation in a-e.

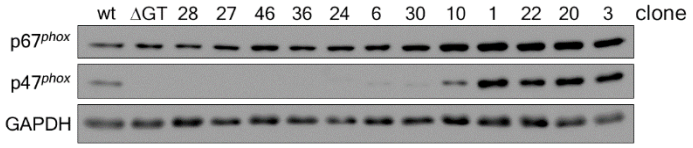
Collectively, this set of experiments demonstrates that around 60% of CRISPR-Cas9-treated PLB-985 *NCF1* ΔGT cells displayed correction of ΔGT mutation in mutated *NCF1* and its pseudogenes, together with restoration of $p47^{phox}$ protein expression and NADPH oxidase function.
Observed levels of ROS generation among the clones were diverse, and not correlated to the number of corrected *NCF1* loci or the level of $p47^{phox}$ protein expression. The reason for this is likely the individual genetic configuration of CRISPR-Cas9-edited *NCF1* loci, as indel mutations can partially rescue NADPH oxidase activity (**Figure S1)**. In addition, correction of mutated *NCF1* and/or pseudogene alleles may be diversely regulated on the level of gene transcription and protein translation. Also, CRISPR-Cas9-induced chromosomal rearrangements may contribute to the observed complex genotype/phenotype of tested clones. Notably, a majority of clones that presented large chromosomal deletions (** and ***) did not display genetic correction, $p47^{phox}$ protein expression, or NADPH oxidase activity.

Figure S4



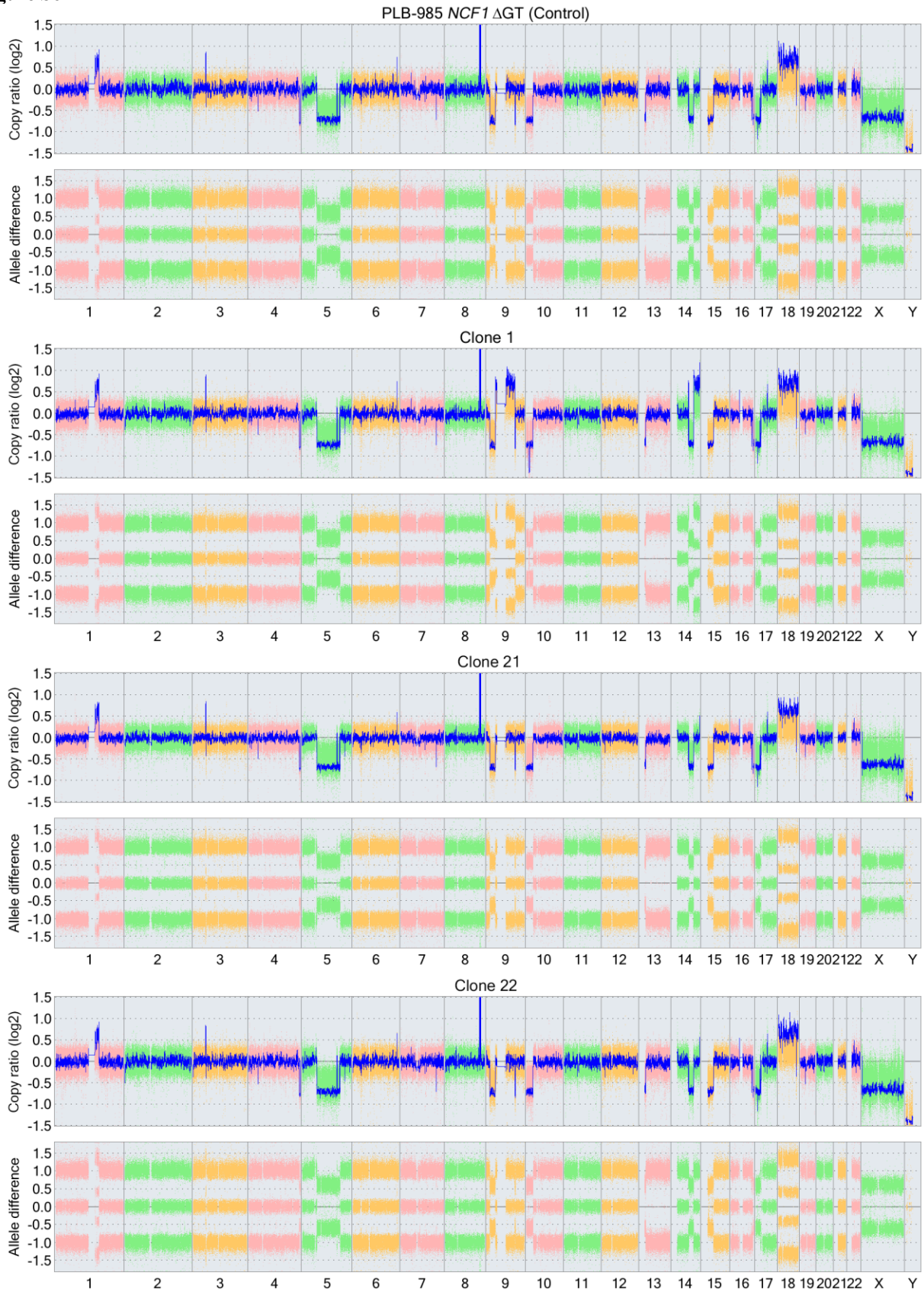
NBT test on individual clones of CRISPR-Cas9-treated PLB-985 *NCF1* ΔGT cell line
Light microscopic images used for the assessment of NADPH oxidase activity in individual clones of CRISPR-Cas9-treated PLB-985 *NCF1* ΔGT cell line after differentiation to granulocytes (NBT score shown in **Figure S3d**). Dark blue precipitates of formazan inside and outside of cells indicate NADPH oxidase-dependent ROS production.

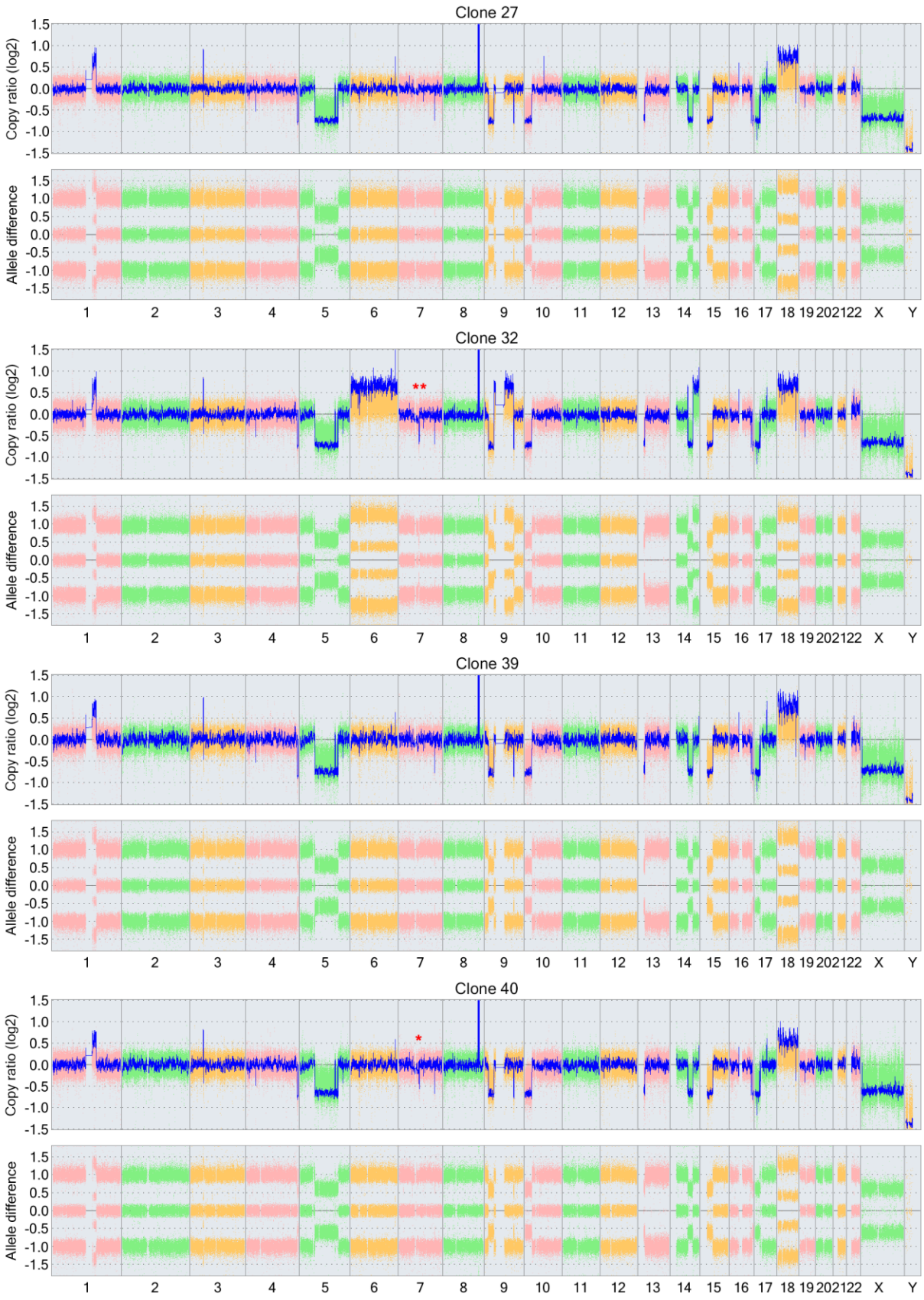
Figure S5

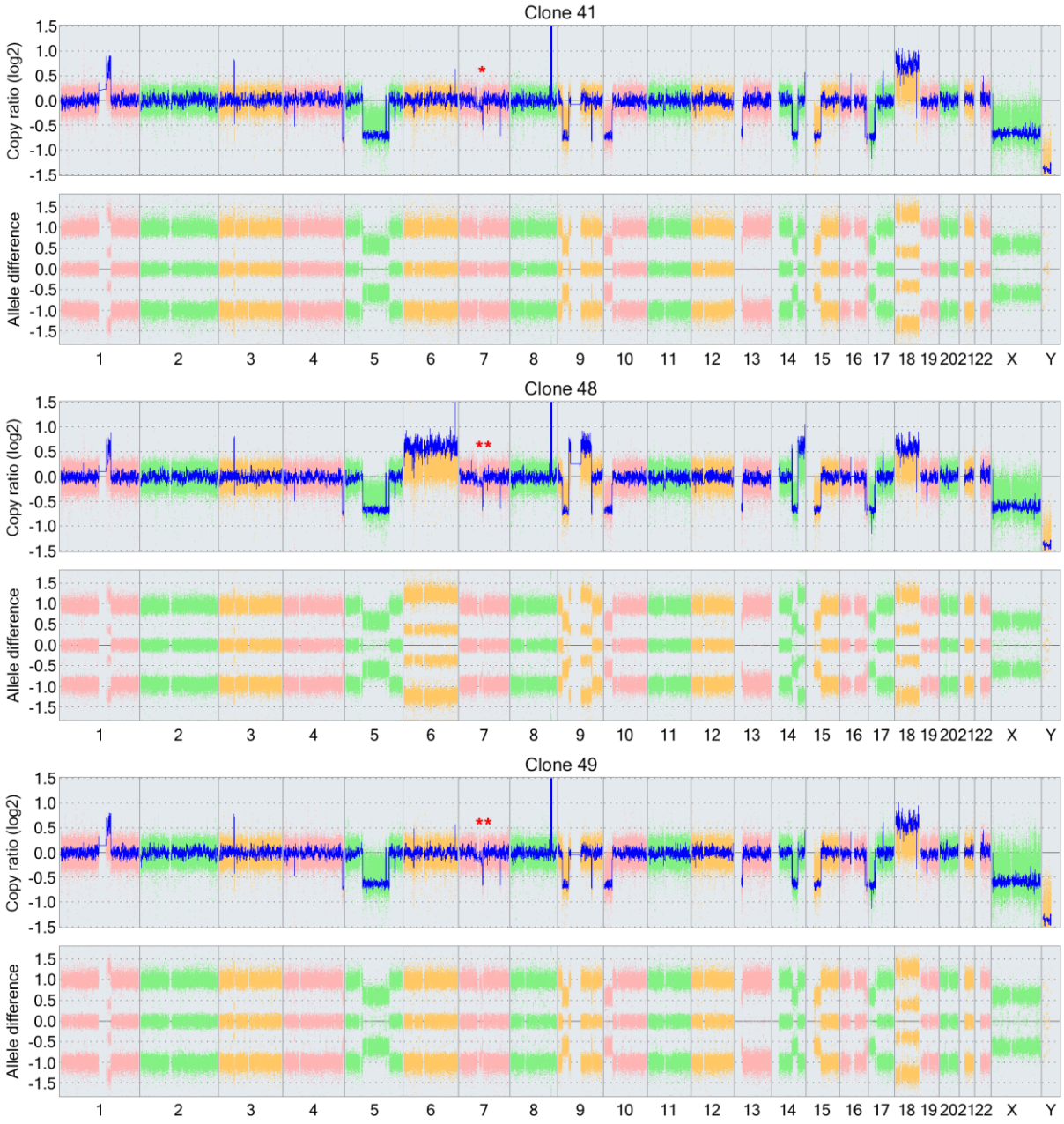


p47*phox* protein expression in selected individual clones of CRISPR-Cas9-treated PLB-985 *NCF1* ΔGT cell line
Western blot of p67*phox*, p47*phox*, and GAPDH in selected clones of CRISPR-Cas9-treated PLB-985 *NCF1* ΔGT cell line after differentiation to granulocytes (**Figure S3e**).

Figure S6


PLB-985 NCF1 ΔGT (Control)
Clone 1
Clone 21
Clone 22
Copy ratio (log2)
Allele difference
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X Y


Clone 27
Copy ratio (log2)
Allele difference
Clone 32
**
Clone 39
Clone 40
*
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X Y



Microarray-based comparative genomic hybridization (aCGH) of selected clones of CRISPR-Cas9-treated PLB-985 *NCF1* ΔGT cell line

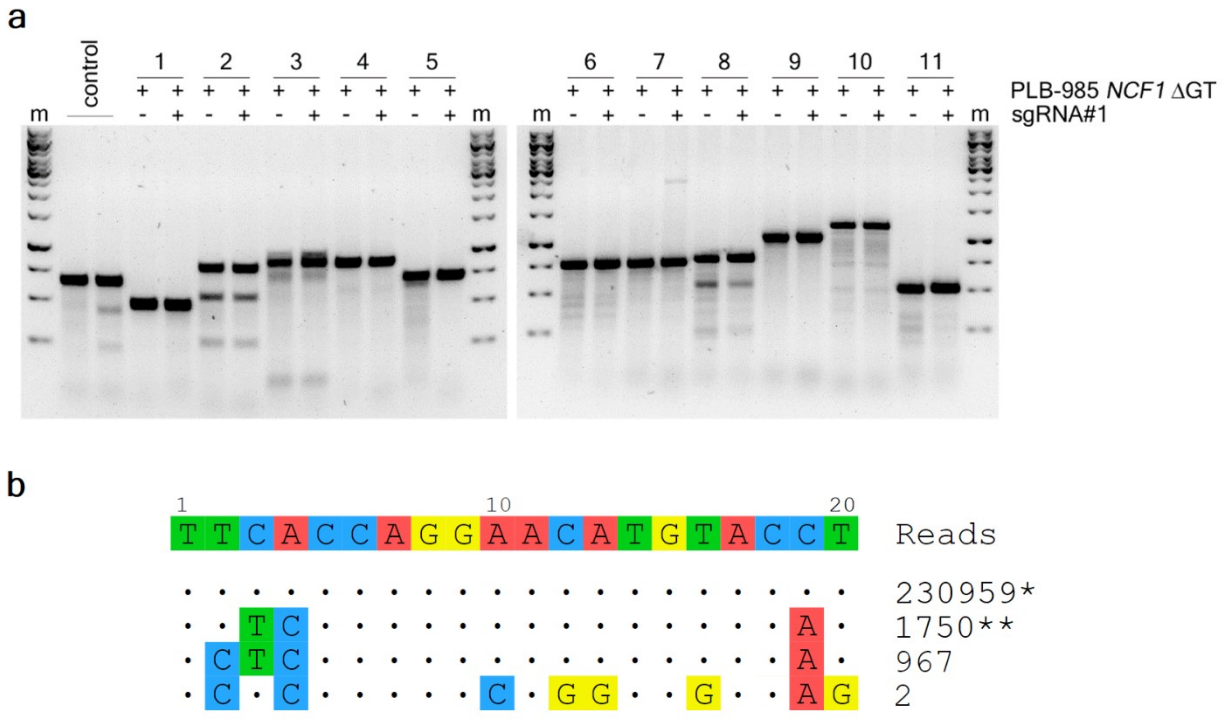
Results are compared to the reference genome using the CytoScan™ HD Array. Copy number ratio (log2) as well as allele differences are shown. Chromosome numbers are listed under the graphs. Aberrations localized between the *NCF1* gene and pseudogene loci (chromosome 7) are marked with either one red asterisk (deletion between *NCF1* and *NCF1C*), or two red asterisks (deletion between *NCF1B* and *NCF1C*).

Identified chromosomal aberrations in the genome of PLB-985 *NCF1* ΔGT cell line are listed in **Table S6**. PLB-985 cell line is a sub clone of the HL-60 cell line,[48] and the aCGH array readout for PLB-985 *NCF1* ΔGT cell line retained aberrations characteristic for the parental cell line, including trisomy of chromosome 18 and monosomy of the X chromosome.

De novo alternations found in selected CRISPR-Cas9-treated clones are listed in **Table S7**. Inspection of the break points of duplicated or deleted chromosomal fragments, 20 000 bp window around the break point, confirmed the presence of the targeted sequence within *NCF1* loci located on chromosome 7. Deletions found between these loci most likely result from CRISPR-Cas9-mediated simultaneous induction of DSBs at the on-target sites. On the other

hand, other chromosomal aberrations could not be attributed to off-target CRISPR-Cas9 activity, as there were no sequences identified within the break points with fewer mismatches than 6 nucleotides.

Figure S7



T7 Endonuclease I assay and GUIDE-seq off-target detection

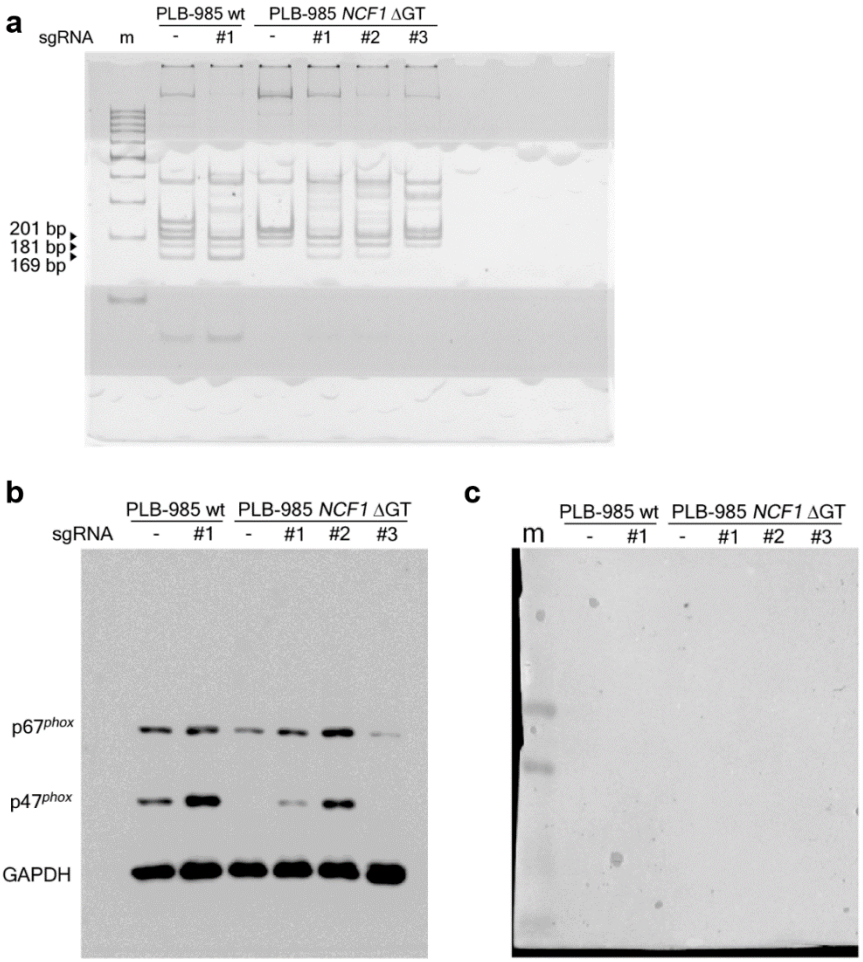
(**a**) The potential off-target sites were predicted with COSMID (CRISPR Off-target Sites with Mismatches, Insertions, and Deletions) analysis tool (https://crispr.bme.gatech.edu/; see **Table S2**).[49] The regions surrounding these sites were PCR amplified from genomic DNA of PLB-985 *NCF1* ΔGT and the bulk culture of CRISPR-Cas9-treated PLB-985 *NCF1* ΔGT cells (sgRNA #1). Corresponding PCR products were mixed in a 1:1 ratio, while PCR amplification products from PLB-985 *NCF1* ΔGT were used as controls along tested samples (total of 200 ng DNA). The lanes marked as 'control' contain PCR products of Control C (lane 2) as well as 1:1 mixture of PCR products of Control C and Control G (lane 3) of the Surveyor® Mutation Detection Kit For Standard Gel Electrophoresis (Integrated DNA Technologies, Leuven, Belgium). The samples were denatured at 95 °C for 5 minutes, slowly renatured, and digested using T7 Endonuclease I (New England Biolabs) according to the manufacturer's instructions.

None of the tested potential off-target sites exhibited additional cleavage, indicating that sgRNA #1 specifically guides Cas9 towards the mutated *NCF1* as well as *NCF1* pseudogenes in PLB-985 *NCF1* ΔGT.

(**b**) PLB-985 *NCF1* ΔGT cells have been nucleofected with a Cas9, GFP and sgRNA #1 encoding plasmid, as described the Methods, in the presence of 10 pmol of Genome-wide Unbiased Identification of DSBs Enabled by Sequencing (GUIDE-seq) dsDNA oligonucleotide. GFP expressing cells were sorted, expanded and subjected to GUIDE-seq analysis.[24] Sequencing was performed using an Illumina MiSeq sequencer (Illumina Inc., San Diego. CA, USA). MiSeq paired-reads were analyzed using guideseq (v1.1b5, mapq = 0) against the human reference genome (Ensembl GRCh38.p10) for identification of off-targets.

The on-target site represents combined reads mapped to all *NCF1* loci (*). Three off-target sequences have been identified by GUIDE-seq, however, the most frequently cleaved off-target (**; 0.76% of the cleavage at the on-target) is represented within five distinct genomic locations (**Table S3**).

Figure S8



Uncropped and/or unexposed gel and membrane shown in Figure 1 of the main text

(**a**) An image of the uncropped PAGE gel of PCR-RFLP shown in **Figure 1C**.

(**b**) An image of the uncropped western blot membrane displaying expression of p67^{phox} (top), p47^{phox} (middle), and GAPDH (bottom), shown in **Figure 1F**.

(**c**) An image of the unexposed western blot membrane shown in b displays the lane for a protein ladder 'm' (PageRuler™ Plus Prestained Protein Ladder 10 to 250 kDa, Thermo Fisher Scientific).

Supplemental Tables

Table S1

Known genes with disease-causing pseudogenes, co-localized on the same chromosome.[6,11,36-38,50-55]

Gene	Disease	OMIM ID	Genomic location	Coordinates (GRCh38)	Pubmed ID
ABCC6	Pseudoxanthoma elasticum (PXE)	603234	16p13.11	16:16,149,564-16,223,616	11474653
CRYBB2	Autosomal dominant cataract	123620	22q11.23	22:25,211,659-25,231,868	3436525
CYP21A2	Congenital adrenal hyperplasia	613815	6p21.33	6:32,038,315-32,041,669	3486422
FOLR1	Neural tube defects	136430	11q13.4	11:72,189,557-72,196,322	1717147
GBA	Type 2 Gaucher disease	606463	1q22	1:155,234,447-155,244,861	2914709
IDS	Hunter syndrome	300823	Xq28	X:149,476,989-149,505,353	7633410
IGLL1	B cell deficiency and agammaglobulinemia	146770	22q11.23	22:23,573,124-23,580,547	12384776
NCF1	Chronic granulomatous disease	608512	7q11.23	7:74,773,961-74,789,375	9329953
PKD1	Autosomal dominant polycystic kidney disease	601313	16p13.3	16:2,088,707-2,135,897	11414761
SBDS	Shwachman-Bodian-Diamond syndrome	607444	7q11.21	7:66,987,676-66,995,695	12496757
RP9	Autosomal dominant retinitis pigmentosa	607331	7p14.3	7:33,094,796-33,109,389	16671097

Table S2

Potential off-target sites predicted by COSMID Off-Target Analysis Tool.[49]

	Off-target sequence	Mismatch	Cut site	Score	Gene	Primer Forward	Primer Reverse	PCR [bp]	Position
sgRNA#1	TTCACCAGGAACATGTACCTNGG	-	-	-	-	-	-	-	-
1	**G**TCACCA**A**GAACATGTAC**A**TGGG	3	76852825	5.39	*ZFHX4*	AGTGAGTGCAGACTGCCAAA	AAGCTTGTCATAGGCCTCCC	454	Chr8:76852444-76852897
2	T**CA**ACCAGGAACATGTA**A**CTTGG	3	19965860	4.28	*LPL*	TAAGGCTCCTTCATGTGGCG	CCAGGAACCTCTCCACCCTT	766	Chr8:19965618-19966383
3	T**C**CACCA**CT**AACATGTACCTAGG	3	120164679	0.75	*SEC22B*	GCACCCAGCCTATTAGCCGT	ACCCAGCAAACCCATTAATGGGT	929	Chr1:120163868-120164796
4	T**C**CACCAGGA**G**CATGT**T**CCTGGG	3	30276197	3.83	*AL021921.1*	GACAGCCAGACATGGATAACCG	GCCACGCAGTGAGTAAGCTG	828	Chr1:30276795-30275968
5	TTC**T**CCA**T**GAACATGT**T**CCTTGG	3	129559878	3.44	*TMEM132D*	GCTTGGTGCCCACCTATGAA	CACAAATGGGAAGACGCTGG	686	Chr12:129560157-129559472
6	TTC**T**CCAGG**GG**CATGTACCTAGG	3	102323741	1.37	*ZNF839*	ACTCGGGATCACTACTGGCA	TCTCCTGCAACTTGAAAGGGC	740	Chr14:102323227-102323966
7	TTCA**T**CAGGA**T**CATGT**T**CCTAGG	3	15897717	3.89	*LOC102724957*	CAGACAAGGCAGGAGGCAG	CTCCAGCAGTCAAGATCTAGCAC	731	Chr11:15898188-15897458
8	TTCAC**A**AGGA**T**CATGTA**T**CTAGG	3	41091359	4.91	*CNTN1*	CATAGTGGCCTGAAGGCAGTAT	TGGCACTTGAGACACTATTGATCC	769	Chr12:41091848-41091080
9	TTCACC**T**G**T**AACATG**G**ACCTGGG	3	4084965	2.88	*C1orf174*	GAAGAGCAGAGTGGAGGCTC	AGCTGTCTGTCGCACTGGAC	1017	Chr1:4084575-4085591
10	TTCACCAG**CT**ACATGTACCTGGG	2	14512241	0.85	*NFIB*	GGCTAGCCTGAGATGACACAG	GAGAAAAGAAAGACGGGAGGTAGG	1204	Chr9:14511591-14512794
11	TTCACCAGG**GG**CATGT**C**CCTGGG	3	139751840	4.20	*NXPH2*	CAGTCCTGGGGCAAGAACA	CTAGGCTGGATACAGTGACTGC	493	Chr2:139751463-139751955

Table S3

Off-target sites detected by GUIDE-seq.[24] The shaded entries represent one off-target sequence repeated five times in the reference genome.

	Off-target sequence	Position	Strand	Mismatch	Reads	Gene or nearest genes
sgRNA#1	TTCACCAGGAACATGTACCTNGG	-	+/-	-	230959	-
1	TT**TC**CCAGGAACATGTAC**A**TGGG	Chr14:18601333-18601355	-	3		*OR11H12*
2	TT**TC**CCAGGAACATGTAC**A**TGGG	Chr14:19713678-19713700	+	3		*OR11H2*
3	TT**TC**CCAGGAACATGTAC**A**TGGG	Chr15:21574969-21574991	+	3	1750	*POTEB3; OR4N4*
4	TT**TC**CCAGGAACATGTAC**A**TGGG	Chr15:22010291-22010313	+	3		*POTEB; OR4M2*
5	TT**TC**CCAGGAACATGTAC**A**TGGG	Chr22:15528375-15528397	-	3		*OR11H1*
6	T**CTC**CCAGGAACATGTAC**A**TGGG	Chr14:19180507-19180529	-	4	967	*POTEM; POTEG*
7	T**C**C**C**CCAGG**C**A**GG**TG**G**AC**AG**GGG	Chr14:104687804-104687826	+	8	2	*TMEM179; INF2*

Table S4
A list of indel mutations at the CRISPR-Cas9 on-target cleavage site identified by SMRT sequencing within CRISPR-Cas9-treated PLB-985 *NCF1* ΔGT cells (sgRNA #1).

Reference sequences (*NCF1: GTGT*; mutated *NCF1*, *NCF1B*, *NCF1C*: ΔGT)	
GGTCCCCCGACTCTGGCTTTCCCCCAGGTGTACATGTTCCTGGTGAAATGGCAGGACCTGTCGGAG	GTGT
GGTCCCCCGACTCTGGCTTTCCCCCAGGT--ACATGTTCCTGGTGAAATGGCAGGACCTGTCGGAG	ΔGT
Identified indels	SMRT reads [%]
GGTCCCCCGACTCTGGCTTTCCCCCAGG---ACATGTTCCTGGTGAAATGGCAGGACCTGTCGGAG	16.4
GGTCCCCCGACTCTGGCTTTCCCCCAGGT-TACATGTTCCTGGTGAAATGGCAGGACCTGTCGGAG	8.6
GGTCCCCCGACTCTGGCTTTCCCCCA----TACATGTTCCTGGTGAAATGGCAGGACCTGTCGGAG	7.1
GGTCCCCCGACTCTGGCTTTCCCCCAGGTG---------------AAATGGCAGGACCTGTCGGAG	5.3
GGTCCCCCGACTCTGGCTTTCCCCCAGG-------------GGTGAAATGGCAGGACCTGTCGGAG	4.5
GGTCCCCCGACTCTGGCTTTCCCCCAGGT--------TCCTGGTGAAATGGCAGGACCTGTCGGAG	4.7
GGTCCCCCGACTCTGGCTTAC------------ATGTTCCTGGTGAAATGGCAGGACCTGTCGGAG	3.8
GGTCCCCCGACTCTGGCTTTCCCCCA--------TGTTCCTGGTGAAATGGCAGGACCTGTCGGAG	2.5
GGTCCCCCGACTCTGGCTTTCCCCCAGGT---------CCTGGTGAAATGGCAGGACCTGTCGGAG	1.6
GGTCCCCCGACTCTGGCTTTCCCCCAGGGAGGTGTACATGTTCCTGGTGAAATGGCAGGACCTGTCGGAG	1.4
GGTCCCCCGACTCTGGCTTTCCCCCAG---TACATGTTCCTGGTGAAATGGCAGGACCTGTCGGAG	1.4
GGTCCCCCGACTCTGGCTTTCCCCCAGGTTGTACATGTTCCTGGTGAAATGGCAGGACCTGTCGGAG	1.4
GGTCCCCCGACTCTGGCTTTCCCCCAGGT----ATGTTCCTGGTGAAATGGCAGGACCTGTCGGAG	1.4
GGTCCCCCGACTCTGGCTTTCCCCCAGGT------------GGTGAAATGGCAGGACCTGTCGGAG	1.2
GGTCCCCCGACTCTGGCTTTCCCCCAGGTTTGTACATGTTCCTGGTGAAATGGCAGGACCTGTCGGAG	1.2
GGTCCCCCGACTCTGGCTTTCCCCCAGG----CATGTTCCTGGTGAAATGGCAGGACCTGTCGGAG	1.0
GGTCCCCCGACTCTGGCTTTCCCCCAGG--------------GTGAAATGGCAGGACCTGTCGGAG	1.0
GGTCCCCCGACTCTGGCTTTCCCCCAG-T-TACATGTTCCTGGTGAAATGGCAGGACCTGTCGGAG	0.9
GGTCCCCCGACTCTGG---------------------------TGAAATGGCAGGACCTGTCGGAG	0.8
GGTCCCCCGACTCTGGCTTTCCCCCAG----ACATGTTCCTGGTGAAATGGCAGGACCTGTCGGAG	0.8
GGTCCCCCGACTCTGGCTTTCCCCCAGGGGTACATGTTCCTGGTGAAATGGCAGGACCTGTCGGAG	0.7
GGTCCCCCGACTCTGGCTTTCCCCCAGGGTGTACATGTTCCTGGTGAAATGGCAGGACCTGTCGGAG	0.7
GGTCCCCCGACTCTGGCTTTCCCCCAGG-GTACATGTTCCTGGTGAAATGGCAGGACCTGTCGGAG	0.7
GGTCCCCCGACTCTGGCTTTCCCCCAGG-------GTTCCTGGTGAAATGGCAGGACCTGTCGGAG	0.7
GGTCCCCCGACTCTGGCTTTCCCCCAGGGGTGTACATGTTCCTGGTGAAATGGCAGGACCTGTCGGAG	0.6
GGTCCCCCGACTCTGGCTTTCCCCCAGGT-------TTCCTGGTGAAATGGCAGGACCTGTCGGAG	0.6
GGTCCCCCGACTCTGGCTTTCC------------TGTTCCTGGTGAAATGGCAGGACCTGTCGGAG	0.5
GGTCCCCCGACTCTGGCTTTCCCCCAGG----------CCTGGTGAAATGGCAGGACCTGTCGGAG	0.5
GGTCCCCCGACTCTGGCTTTCCCCCA--------TGT-CCTGGTGAAATGGCAGGACCTGTCGGAG	0.5
GGTCCCCCGACTCTGGCTTT----------TACATGTTCCTGGTGAAATGGCAGGACCTGTCGGAG	0.5
GGTCCCCCGACTCTGGCTTTCCCCC---TG--CATGTTCCTGGTGAAATGGCAGGACCTGTCGGAG	0.4
GGTCCCCCGACTCTGGCTTTCCCCCAGG----------------GGAATGGCAGGACCTGTCGGAG	0.3
GGTCCCCCGACTCTGGCTTTCCCCCAGG-G-ACATGTTCCTGGTGAAATGGCAGGACCTGTCGGAG	0.3
GGTCCCCCGACTCTGGCTTTCCCCCAGG-G-------TCCTGGTGAAATGGCAGGACCTGTCGGAG	0.3
GGTCCCCCGACTCTGGCTTTCCCCCAGG-----ATGTT---GGTGAAATGGCAGGACCTGTCGGAG	0.3
GGTCCCCCGACTCTGGCTTTCCCCCAGGTGT------TCCTGGTGAAATGGCAGGACCTGTCGGAG	0.3
GGTCCCCCGACTCTGGCTTTCCCCCA--------TGTT---GGTGAAATGGCAGGACCTGTCGGAG	0.3

GGTCCCCCGACTCTGGCTTTCCCCCAGG-----ATGTTC-TGGTGAAATGGCAGGACCTGTCGGAG	0.3
GGTCCCCCGACTCTGGCTTTCCCCCAGG----------AAATGGCAGGACCTGTCGGAG	0.3
GGTCCCCCGACTCTGGCTTTCCC--------ACATGTTCCTGGTGAAATGGCAGGACCTGTCGGAG	0.3
GGTCCCCCGACTCTGGCTTTCCCCCAGGT-TACATGTTCCTGGTGAAATGGCAGGACCTGTCGGAG	0.3
GGTCCCCCGACTCTGGCTTTCCCC---------ATGTTCCTGGTGAAATGGCAGGACCTGTCGGAG	0.2
GGTCCCCCGACTCTGGCTTCCCC--AGGT--ACATGTTCCTGGTGAAATGGCAGGACCTGTCGGAG	0.2
GGTCCCCCGACTCTGGCTTTCCCCCAGGGGGTACATGTTCCTGGTGAAATGGCAGGACCTGTCGGAG	0.2
GGTCCCCCGACTCTGGCTTTCCCC-------ACATGTTCCTGGTGAAATGGCAGGACCTGTCGGAG	0.2
GGTCCCCCGACTCTGGCTTTCCCCCAGGT--ACA--TTCCTGGTGAAATGGCAGGACCTGTCGGAG	0.2

Table S5

Primers used for the assessment of CNV by qPCR (**Figure2A and Figure 2B**).

Gene	**Location relative to *NCF1* loci**	**Forward primer**	**Reverse primer**
CALN1	Centromeric side of *NCF1B*	GGTGATTGGCTGTGTCTTCC	CCGGCTAAGTAATCAGCTCCA
EIF4H	Between *NCF1B* and *NCF1*	TCAGAAAAGGTGGACCAGATGAC	GGAACGTTCACGAGTTGGAAT
WBSCR16	Between *NCF1* and *NCF1C*	CCTGGGATTCCATCTGGAGC	CGATTCTCTTCCTGAGGGGC
HIP1	Telomeric side of *NCF1*	GAAGCCTTGCCTCCTCAACT	GGTCTGGTTGGAGATGGGTG
SOD1	Chromosome 21 (reference)	CAGAGGCCTTGGGACATAGC	ATGGGGCTGCACCTGATTTC

Table S6
Chromosomal aberrations identified for the PLB-985 *NCF1* ΔGT cell line.

	Chromosome	Locus	Start	End	Size [kb]	Mean log2 ratio	Median log2 ratio	Copy number
1	1	q21.1:q23.2	145382123	159863470	14481.3	0.2790	0.2892	3
2	3	p21.31	45599905	47567580	1967.7	0.3130	0.3204	3
3	3	p21.31	47570543	47654374	83.8	-0.5346	-0.5187	1
4	4	p15.1	34779031	34824169	45.1	-2.0251	-2.0282	0
5	4	q28.3	135061532	135094781	33.2	-0.5964	-0.5888	1
6	4	q35.1:q35.2	186676232	188997015	2320.8	-0.6406	-0.6358	1
7	5	q11.2:q.23.3	53675286	127682183	74006.9	-0.5359	-0.5318	1
8	5	q23.3:q31.1	128172849	135619529	7446.7	-0.5231	-0.5175	1
9	5	q31.1:q31.3	135842660	139578983	3736.3	-0.5387	-0.5366	1
10	5	q33.3	156811544	156903327	91.8	-0.5009	-0.5014	1
11	6	q13	74590396	74601723	11.3	-1.5305	-1.5288	0
12	6	q26	162722580	162912831	190.3	0.3081	0.3003	3
13	7	q32.2	130810971	131005889	194.9	-0.5437	-0.5412	1
14	8	q24.13	126224383	126547568	323.2	1.1117	1.1269	4
15	8	q24.13:q24.21	126712969	127390069	677.1	1.1220	1.1502	4
16	8	q24.21	128075476	128345268	269.8	1.1384	1.1348	4
17	8	q24.21	128690529	128771759	81.2	1.2554	1.2835	4
18	8	q24.21	129987555	130209520	222.0	1.4051	1.4172	4
19	8	q24.21	130366403	130697499	331.1	1.4732	1.4976	4
20	9	p23:p21.1	11774820	32397138	20622.3	-0.5663	-0.5651	1
21	9	q31.1	104636449	106946358	2309.9	-0.6005	-0.5927	1
22	10	p15.3:p12.1	100026	25425701	25325.7	-0.5498	-0.5464	1
23	13	q11:q12.12	19436286	23397468	3961.2	-0.5003	-0.5002	1
24	14	q23.2	62173186	62630023	456.8	-0.5433	-0.5439	1
25	14	q23.2:q31.1	64716201	81909742	17193.5	-0.5285	-0.5254	1
26	14	q32.33	106329183	106723341	394.2	0.3315	0.3058	3
27	15	q11.2:q15.3	22770421	43976996	21206.6	-0.5398	-0.5348	1
28	16	q23.2:q23.3	81280004	81897623	617.6	-0.5104	-0.5030	1
29	16	q24.1:q24.3	85510923	90146767	4635.8	-0.5154	-0.5048	1
30	17	p13.3:p11.2	525	17099612	17099.1	-0.5364	-0.5257	1
31	17	p11.2	18144526	20753871	2609.3	-0.5252	-0.5234	1
32	17	q12	36350597	36404135	53.5	-0.5300	0.5234	1
33	18	trisomy				0.3026	0.3080	
34	19	p12	20598429	20720704	122.3	-0.486	-0.447	1
35	22	q11.23:q12.1	25656237	25922333	266.1	0.4080	0.4057	3
36	X	monosomy				-0.5191	-0.5161	

Table S7
De novo chromosomal aberrations identified for selected clones of CRISPR-Cas9-treated PLB-985 *NCF1* ΔGT cell line by aCGH.

clone 1								
	Chromosome	Locus	Start	End	Size [kb]	Mean log2 ratio	Median log2 ratio	Copy number
1	9	p21.1:q31.1	32406181	104628430	72222.2	0.3072	0.3126	3
2	10	p14	9668766	11590968	1922.2	-2.0910	-2.0904	0
3	13	q21.2	61021403	61040152	18.7	-0.4828	-0.5300	1
4	14	p23.2	62634687	64690561	2055.9	0.2617	0.2653	3
5	14	q31.1:q32.33	81923332	107145067	25221.7	0.3092	0.3146	3
clone 21 No aberrations								
clone 22								
	Chromosome	Locus	Start	End	Size [kb]	Mean log2 ratio	Median log2 ratio	Copy number
1	1	p34.3	35771294	35783945	12.7	-0.5677	-0.5505	1
clone 27								
	Chromosome	Locus	Start	End	Size [kb]	Mean log2 ratio	Median log2 ratio	Copy number
1	10	q22.1	70692429	70932814	240.4	0.318	0.327	3
clone 32								
	Chromosome	Locus	Start	End	Size [kb]	Mean log2 ratio	Median log2 ratio	Copy number
1	6	trisomy				0.2926	0.2980	3
2	7	q11.23	72636883	74141745	1504.9	-0.4664	-0.4736	1
3	7	q11.23	74191290	74578927	387.6	-0.4610	-0.4654	1
4	9	p21.1:q31.1	32406181	104628430	72222.2	0.2918	0.2968	3
5	14	q23.2	62634525	64686542	2052.0	0.2431	0.2469	3
6	14	q31.1:q32.33	81923332	107145067	25221.7	0.2956	0.2999	3
clone 39 No aberrations								
clone 40								
	Chromosome	Locus	Start	End	Size [kb]	Mean log2 ratio	Median log2 ratio	Copy number
1	7	q11.23	74197396	74578927	381,5	-0.332	-0.337	1
clone 41								
	Chromosome	Locus	Start	End	Size [kb]	Mean log2 ratio	Median log2 ratio	Copy number
1	7	q11.23	74191290	74286808	95.5	-0.384	-0.365	1
2	7	q11.23	74404149	74578928	174.8	-0.424	-0.415	1
clone 48								
	Chromosome	Locus	Start	End	Size [kb]	Mean log2 ratio	Median log2 ratio	Copy number
1	6	trisomy				0.2732	0.2778	3

2	7	q11.23	72637752	74141492	1503.7	-0.4462	-0.4467	1
3	7	q11.23	74190888	74578927	388.0	-0.4162	-0.4404	1
4	9	p21.1:q31.1	32409908	104622867	72213.0	0.2809	0.2837	3
5	14	q23.2	62634753	64704354	2069.6	0.2251	0.2311	3
6	14	q31.1:q32.33	81939428	107145067	25205.6	0.2830	0.2875	3
clone 49								
	Chromosome	Locus	Start	End	Size [kb]	Mean log2 ratio	Median log2 ratio	Copy number
1	6	p21.33	31566126	31757782	191.7	0.300	0.283	3
2	7	q11.23	72637752	74578927	1941.2	-0.430	-0.436	1

Supplemental references

45. Wohlgensinger, V., Seger, R., Ryan, M. D., Reichenbach, J. & Siler, U. Signed outside: A surface marker system for transgenic cytoplasmic proteins. *Gene Ther.* **17**, 1193–1199 (2010).
46. Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–21 (2012).
47. Pedruzzi, E., Fay, M., Elbim, C., Gaudry, M. & Gougerot-Pocidalo, M.-A. Differentiation of PLB-985 myeloid cells into mature neutrophils, shown by degranulation of terminally differentiated compartments in response to N-formyl peptide and priming of superoxide anion production by granulocyte-macrophage colony-stimulating fact. *Br. J. Haematol.* **117**, 719–26 (2002).
48. Drexler, H. G., Dirks, W. G., Matsuo, Y. & MacLeod, R. A. F. False leukemia-lymphoma cell lines: An update on over 500 cell lines. *Leukemia* **17**, 416–426 (2003).
49. Cradick, T. J., Qiu, P., Lee, C. M., Fine, E. J. & Bao, G. COSMID: A web-based tool for identifying and validating CRISPR/Cas off-target sites. *Mol. Ther. - Nucleic Acids* **3**, e214 (2014).
50. Germain, D. P. Pseudoxanthoma elasticum: evidence for the existence of a pseudogene highly homologous to the ABCC6 gene. *J. Med. Genet.* **38**, 457–61 (2001).
51. Aarts, H. J., Den Dunnen, J. T., Lubsen, N. H. & Schoenmakers, J. G. Linkage between the beta B2 and beta B3 crystallin genes in man and rat: a remnant of an ancient beta-crystallin gene cluster. *Gene* **59**, 127–35 (1987).
52. Higashi, Y., Yoshioka, H., Yamane, M., Gotoh, O. & Fujii-Kuriyama, Y. Complete nucleotide sequence of two steroid 21-hydroxylase genes tandemly arranged in human chromosome: a pseudogene and a genuine gene. *Proc. Natl. Acad. Sci. U. S. A.* **83**, 2841–2845 (1986).
53. Bondeson, M. L. *et al.* Inversion of the IDS gene resulting from recombination with IDS-related sequences in a common cause of the hunter syndrome. *Hum. Mol. Genet.* **4**, 615–621 (1995).
54. Linnebank, M. *et al.* Argininosuccinate lyase (ASL) deficiency: mutation analysis in 27 patients and a completed structure of the human ASL gene. *Hum. Genet.* **111**, 350–9 (2002).
55. Boocock, G. R. B. *et al.* Mutations in SBDS are associated with Shwachman-Diamond syndrome. *Nat. Genet.* **33**, 97–101 (2003).